

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 17:11:26 ; Search time 36 Seconds

(without alignments)  
932.754 Million cell updates/sec

Title: US-09-910-033a-2

Perfect score: 1290

Sequence: 1 MSNRLDGKVAITGTGTLGIG.....NESKATGSEFVVDGVTAA 252

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1290	100.0	252	23	AAO16940
2	1277	99.0	251	18	AAW23407
3	1019	79.0	251	18	AAW23409
4	456	35.3	254	22	AAW49773
5	450	34.9	254	22	AAW47522
6	400	31.0	258	21	AAW56815
7	376	29.1	257	23	ABP38717
8	369	28.6	272	22	AAU37095
9	368	28.5	253	23	ABW54214
10	367.5	28.5	261	13	AAW27757

11	366.5	28.4	261	23	AAE20115	Lactobacillus rham
12	364.5	28.3	261	11	AAW04044	Glucose dehydrogen
13	364.5	28.3	261	13	AAW24018	Thermotable glucos
14	364	28.2	251	22	AAU34193	Staphylococcus aur
15	363.5	28.2	272	21	AAW10740	B. megaterium gluc
16	363.5	28.2	340	21	AAW10741	H. ghilani/B. me
17	360.5	27.9	261	11	AAW03846	Modified glucose d
18	359.5	27.9	261	13	AAW27756	NAD affinity glucos
19	358.5	27.8	261	9	AAW05090	Sequence of glucos
20	358.5	27.8	261	21	AAW96271	B. subtilis glucos
21	358.5	27.8	261	21	AAW54424	Amino acid sequenc
22	356.5	27.6	261	9	AAW80063	Glucose dehydrogen
23	352.5	27.3	262	22	AAW83032	S. epidermidis ope
24	351	27.2	286	22	AAW36287	Pseudomonas aerugi
25	348.5	27.0	263	23	ABP39667	Staphylococcus epi
26	345	26.7	306	22	AAW81644	S. epidermidis ope
27	344	26.7	270	21	AAW42558	Human ORFX ORF2322
28	344	26.7	303	23	ABW26264	Herbically active
29	341	26.4	274	23	ABP39586	Staphylococcus epi
30	340	26.4	270	20	AAW41761	Human PRO474 (UNO5
31	340	26.4	270	21	AAW44317	Human PRO474 (UNO5
32	340	26.4	270	21	AAW24056	Human PRO474 prote
33	340	26.4	270	22	AAW28108	Novel human secret
34	340	26.4	279	22	AAW18296	Human endocrine po
35	336.5	26.1	271	22	AAU37794	Streptococcus pneu
36	336.5	26.1	272	22	AAU38086	Novel human secret
37	334.5	25.9	288	22	AAW28296	Herbically active
38	333.5	25.9	280	23	ABW92449	Streptococcus poly
39	330.5	25.6	254	23	ABP30228	Streptococcus poly
40	330.5	25.6	263	23	ABW28167	Streptococcus poly
41	330.5	25.6	270	23	ABW26580	Secoisolaricresin
42	330.5	25.6	277	21	AAW54415	Listeria monocytog
43	324	25.1	253	23	ABW50010	Ramoplanin biosyn
44	324	25.1	274	23	AAW22161	Maize t52 sequence
45	324	25.1	336	16	AAW06488	

#### ALIGNMENTS

RESULT 1	AAO16940	standard; Protein: 252 AA.
ID	AAO16940	
XX	AAO16940:	
MC	16-MAY-2002 (first entry)	
XX		
DT		
XX		
DE	Recombinant enzyme with increased NAD(H) acceptance.	
XX		
KW	NAD(H) acceptance enzyme; reduced nicotinamide-adenine dinucleotide;	
KM	ketone; alcohol; enantiomerically enriched.	
XX		
OS	Unidentified.	
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 38	/note- "wild-type Gly substituted by Asp"
XX		
PN	EP1176203-A1.	
PD	30-JAN-2002.	
XX		
PF	20-JUN-2001: 2001EP-0114953.	
XX		
PR	27-JUL-2000: 2000DE-1037101.	
XX		
PA	(DEGS ) DECUSSA AG.	
XX		
PI	Riebel B, Hummel W, Bommarilus A;	
XX		
DR	WPI: 2002-173122/23.	

DR N-PSDB: AAL45006.

XX New recombinant enzyme, useful for enantioselective synthesis of e.g.  
PT alcohols, has increased nicotinamide-adenine dinucleotide acceptance as  
PT result of specific mutation -

XX PS Disclosure: Page 7-8; 23pp; German.

CC The present invention relates to a recombinant enzyme with higher NAD(H)  
CC (reduced) nicotinamide-adenine dinucleotide) acceptance than the wild  
CC type, and which has at least one neutral amino acid exchanged for an  
CC acidic amino acid, while retaining the basic amino acid in the co-enzyme  
CC binding site. The enzyme can be used to prepare enantioselectively enriched  
CC organic compounds, especially enantioselective reduction of ketones or  
CC oxidation of alcohols. The present sequence is the enzyme of the  
CC invention.

XX Sequence 252 AA;

Query Match 100.0%; Score 1290; DB 23; Length 252;  
Best Local Similarity 100.0%; Pred. No. 4,3e-117;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNRLDGVVAITITGTLGIGLAIAATKFEVGEAKVWITDRHSDVGEKAAKSVGTPOIOFF 60  
DB 1 MSNRLDGVVAITITGTLGIGLAIAATKFEVGEAKVWITDRHSDVGEKAAKSVGTPOIOFF 60  
QY 61 QHSDSDGWTKLFDATKAGPSTLVNNGIAVNSVEETTTAEMRKLAVNLDGVFF 120  
DB 61 QHSDSDGWTKLFDATKAGPSTLVNNGIAVNSVEETTTAEMRKLAVNLDGVFF 120  
QY 121 GTRIGIORMKKGAGASINNSIEGFVGPDSLCAVNAKCAVRIMSKAALDCAKD 180  
DB 121 GTRIGIORMKKGAGASINNSIEGFVGPDSLCAVNAKCAVRIMSKAALDCAKD 180  
QY 181 VRVTVHPGYIKTPLVVDLPAGEAEMSORRTKTPMGHIGEPNDIAYICVYLAESKFRATG 240  
DB 181 VRVTVHPGYIKTPLVVDLPAGEAEMSORRTKTPMGHIGEPNDIAYICVYLAESKFRATG 240  
QY 241 SEFVVDGTYTAQ 252  
DB 241 SEFVVDGTYTAQ 252

# RESULT 2

AAM23407

ID AAM23407 standard; Protein; 252 AA.

XX AAM23407;

DT 27-MAR-1998 (first entry)

XX Lactobacillus brevis alcohol dehydrogenase.

KW Alcohol dehydrogenase; production; (R)-alcohol; (S)-alcohol.

OS Lactobacillus brevis.

PN EP796914-A2.

PD 24-SEP-1997.

PF 20-MAR-1997; 97EP-0104814.

PR 21-MAR-1996; 96DE-4010984.

PA (BOE) BOEHRINGER MANNHEIM GMBH.

PI Hummel W, Riebel B;

XX WPI; 1997-459831/43.

DR N-PSDB: AAT73132.

XX

PT Lactobacillus brevis alcohol dehydrogenase - useful for production  
PT of optically active alcohol(s)

XX PS Claim 8; Pages 24-26; 34pp; German.

CC The present sequence is a lactobacillus brevis alcohol  
CC dehydrogenase, which retains at least 95% of its activity after 30  
CC minutes at 20-60 degrees C and can be purified to a specific  
CC activity of at least 400 U/mg. The enzyme can be used to produce  
CC (R)-alcohols by enantioselective reduction of ketones of formula  
CC R1-CO-R2, where R1 and R2 = hydrogen (sic) or 1-20C alkyl, alkenyl,  
CC aryl or arylethyl (sic) optionally substituted by halogen, NO2, OH  
CC or 1-20C alkoxy, "an optionally substituted 1-10C alkylene group,  
CC which is substituted by saturated, unsaturated or aromatic  
CC nitrogen, oxygen or sulphur heterocycles, or may be an optionally  
CC substituted polycondensed saturated and/or aromatic group" (sic) in  
CC the presence of the enzyme or cells containing it at 20-60 degrees  
CC C for 0.25-3 hours, or (S)-alcohols by incubating a racemic alcohol  
CC of formula R1-CHOH-R2 in the presence of the enzyme or cells  
CC containing it at 20-60 degrees C for 0.25-3 hours.

XX Sequence 252 AA;

Query Match 99.0%; Score 1277; DB 18; Length 252;  
Best Local Similarity 99.2%; Pred. No. 7,9e-116;  
Matches 250; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSNRLDGVVAITITGTLGIGLAIAATKFEVGEAKVWITDRHSDVGEKAAKSVGTPOIOFF 60  
DB 1 MSNRLDGVVAITITGTLGIGLAIAATKFEVGEAKVWITDRHSDVGEKAAKSVGTPOIOFF 60  
QY 61 QHSDSDGWTKLFDATKAGPSTLVNNGIAVNSVEETTTAEMRKLAVNLDGVFF 120  
DB 61 QHSDSDGWTKLFDATKAGPSTLVNNGIAVNSVEETTTAEMRKLAVNLDGVFF 120  
QY 121 GTRIGIORMKKGAGASINNSIEGFVGPDSLCAVNAKCAVRIMSKAALDCAKD 180  
DB 121 GTRIGIORMKKGAGASINNSIEGFVGPDSLCAVNAKCAVRIMSKAALDCAKD 180  
QY 181 VRVTVHPGYIKTPLVVDLPAGEAEMSORRTKTPMGHIGEPNDIAYICVYLAESKFRATG 240  
DB 181 VRVTVHPGYIKTPLVVDLPAGEAEMSORRTKTPMGHIGEPNDIAYICVYLAESKFRATG 240  
QY 241 SEFVVDGTYTAQ 252  
DB 241 SEFVVDGTYTAQ 252

# RESULT 3

AAM23409

ID AAM23409 standard; Protein; 251 AA.

XX AAM23409;

DT 27-MAR-1998 (first entry)

XX Lactobacillus kefir alcohol dehydrogenase.

KW Alcohol dehydrogenase; production; (R)-alcohol; (S)-alcohol.

OS Lactobacillus kefir.

PN Key Location/Qualifiers

FT Misc-difference 46 /note= "not defined in specification"

FT Misc-difference 47 /note= "not defined in specification"

FT Misc-difference 48 /note= "not defined in specification"

FT Misc-difference 49 /note= "not defined in specification"

FT Misc-difference 50 /note= "not defined in specification"

FT Misc-difference 51 /note- "not defined in specification"  
 FT Misc-difference 52 /note- "not defined in specification"  
 FT Misc-difference 53 /note- "not defined in specification"  
 FT Misc-difference 54 /note- "not defined in specification"  
 FT Misc-difference 55 /note- "not defined in specification"  
 FT Misc-difference 56 /note- "not defined in specification"  
 FT Misc-difference 57 /note- "not defined in specification"  
 FT Misc-difference 58 /note- "not defined in specification"  
 FT Misc-difference 59 /note- "not defined in specification"  
 FT Misc-difference 60 /note- "not defined in specification"  
 FT Misc-difference 61 /note- "not defined in specification"  
 FT Misc-difference 62 /note- "not defined in specification"  
 FT Misc-difference 63 /note- "not defined in specification"  
 FT Misc-difference 64 /note- "not defined in specification"  
 FT Misc-difference 65 /note- "not defined in specification"  
 FT Misc-difference 66 /note- "not defined in specification"  
 FT Misc-difference 67 /note- "not defined in specification"  
 FT Misc-difference 68 /note- "not defined in specification"  
 FT Misc-difference 69 /note- "not defined in specification"  
 FT Misc-difference 70 /note- "not defined in specification"  
 FT Misc-difference 71 /note- "not defined in specification"  
 FT Misc-difference 130 /note- "not defined in specification"  
 FT Misc-difference 131 /note- "not defined in specification"  
 FT Misc-difference 132 /note- "not defined in specification"  
 FT Misc-difference 133 /note- "not defined in specification"  
 XX EP796914-A2.  
 XX  
 XX  
 PD 24-SEP-1997.  
 XX  
 PF 20-MAR-1997: 97EP-0104814.  
 XX  
 PR 21-MAR-1996: 96DE-4010984.  
 XX  
 XX (BOEF ) BOEHRINGER MANNHEIM GMBH.  
 XX  
 XX Hummel W, Riebel B;  
 XX  
 DR WPI: 1997-459831/43.  
 XX  
 XX  
 PT Lactobacillus brevis alcohol dehydrogenase - useful for production  
 of optically active alcohol(s)  
 XX  
 XX Example 8: Pages 28-30; 34pp; German.  
 XX  
 CC The present Lactobacillus kefir alcohol dehydrogenase (ADH) was  
 used in the isolation of a L. brevis ADH, which retains at least  
 93% of its activity after 30 minutes at 20-60 degrees C and can be  
 purified to a specific activity of at least 400 U/mg. The enzyme  
 can be used to produce (R)-alcohols by enantioselective reduction  
 of ketones of formula R1-CO-R2, where R1 and R2 = hydrogen (sic)  
 or 1-20C alkyl, alkenyl, aryl or arylenyl (sic) optionally

CC substituted by halogen, NO2, OH or 1-20C alkoxy, "an optionally  
 CC substituted 1-10C alkylene group, which is substituted by  
 CC saturated, unsaturated or aromatic nitrogen, oxygen or sulphur  
 CC heterocycles, or may be an optionally substituted polycondensed  
 CC saturated and/or aromatic group" (sic) in the presence of the  
 CC enzyme or cells containing it at 20-60 degrees C for 0.25-3 hours,  
 CC or (S)-alcohols by incubating it at 20-60 degrees C for 0.25-3 hours,  
 CC R1-CHOH-R2 in the presence of the enzyme or cells containing it at  
 CC 20-60 degrees C for 0.25-3 hours.  
 XX  
 SO Sequence 251 AA:  
 Query Match 79.0%; Score 1019; DB 18; Length 251;  
 Best Local Similarity 80.5%; Pred. No. 9,3e-91;  
 Matches 202; Conservative 14; Mismatches 35; Indels 0; Gaps 0;  
 OY 2 SNRLDKVAILITGGTIGIAIAIFKFEAGAKWITRHSVDGGRKAKSVTPPQIQFQ 61  
 DB 1 TDRLKGRVAILVTGGTIGIALADKFEAGAKVITGRHADVGRKXXXXXXXXXXXXX 60  
 OY 62 HDSSDEGWTLPDATEKAGPVTLVNAGIAVNSKEETTAEMRKLAVINDGVFG 121  
 DB 61 XXXXXXXXXXXXLPDATEKAGPVTTVNAGIAVNSKEETTAEMRKLAVINDGVFG 120  
 OY 122 TRLGIRMKKNGIGASTINNSIEGVDPSLGAVNASKAVIRMSAALDALKDYDV 181  
 DB 121 TRLGIRMKKNGIGASTINNSIEGVDPSLGAVNASKAVIRMSAALDALKDYDV 180  
 OY 182 RVNTHVPGYIKTPVNDLPCEAFEMSORTPKMHCEPNDIAICYVLANSKEKPTGS 241  
 DB 181 RVNTHVPGYIKTPVNDLPCEAFEMSORTPKMHCEPNDIAICYVLANSKEKPTGS 240  
 OY 242 EFVVDGSGYTQ 252  
 DB 241 EFVVDGSGYTQ 251  
 RESULT 4  
 AAB49773  
 ID AAB49773 standard; protein; 254 AA.  
 XX  
 AC AAB49773;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Protein with acetylpyridine derivative reducing action.  
 XX  
 KM Optically-active pyridineethanol derivative; asymmetric reduction.  
 XX  
 OS Candida maris.  
 XX  
 PN WO200105996-A1.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 28-JUN-2000: 2000MO-JP04237.  
 XX  
 PR 21-JUL-1999: 99JP-0206503.  
 XX  
 XX (KANF ) KANEKA CORP.  
 XX  
 XX Kawano S, Horikawa M, Yasohara Y, Hasegawa J;  
 DR  
 DR WPI: 2001-159546/16.  
 DR N-PSDB; AAF29375, AAF29376.  
 XX  
 PT Efficient, high-yielding preparation of optically-active  
 PT pyridineethanol derivatives by stereoselectively reducing  
 PT acetylpyridine derivatives e.g. with enzyme having asymmetric reduction  
 activity, for pharmaceutical intermediates -  
 XX  
 PS Claim 14; Fig 1; 76pp; Japanese.

CC This invention relates to a process for producing optically-active  
 CC pyridineethanol derivatives by stereoselectively reducing acetylpyridine  
 CC derivatives with an enzyme or enzyme source having asymmetric reduction  
 CC activity. The process is efficient and high yielding to give R and S  
 CC isomers by suitable manipulation. The method is for the preparation of  
 CC optically-active pyridineethanol derivatives by stereoselectively  
 CC reducing acetylpyridine derivatives, and also similarly for their  
 CC polycyclic analogues, for use as pharmaceutical and agrochemical  
 CC intermediates, as well as in fine chemical production. The present  
 CC sequence represents the amino acid sequence of the enzyme used in the  
 CC process which has acetylpyridine derivative reducing action.

CC Sequence 254 AA:

Query Match 35.3%; Score 456; DB 22; Length 254;  
 Best Local Similarity 40.2%; Pred. No. 5e-36;  
 Matches 104; Conservative 49; Mismatches 94; Indels 12; Gaps 5;

OY 1 MSNRIDGKVAITITGTLGIGLAATKFEVBEKAVMITD-----RHSVGEKAKSVGTP 54  
 DB 1 MSYFNANKVLTGTLGIGLAATKFEVBEKAVMITD-----RHSVGEKAKSVGTP 60  
 OY 55 DOIQFQHDSDSDGWTKLPDATEKAFGPVSTLVNNAIGAVNKSVEETTAEKRLAVN 114  
 DB 61 N-VQYIADASKADENKELISSETLSAFGLDYVCANAGIATFTQTTDISTVWRKYTSIN 119  
 OY 115 LDGVFFGRGLGIQKRNKNGLGASIIINSSIEGFVGDPSLGAVNASKAVRIMSKSALDC 174  
 DB 120 LDGVFMDRLKLAQVFLSKNKPGLAIVMGSISHSYVAFGLSHCAAGKGLKLTQTALEY 179  
 OY 175 ALKDYDVARNVHPGYIKTPLVDDLPGAEEAMSOR-TKTPMGHIGEPNDIATICYTLASN 233  
 DB 180 AAK-GIRVANSVNPGYIKTPLVDDICP--KEHMDVLTQHPGRGKPEETASAVAFICSD 235  
 OY 234 ESKFATGSEFVVDGTYAQ 252  
 DB 236 EATFINGISLVPDGGTYAR 254

RESULT 5

AAB47522 standard; Protein: 254 AA.

AC AAB47522;

DT 04-DEC-2001 (first entry)

DE (R)-2-octanol dehydrogenase.

KM (R)-2-octanol dehydrogenase; ketone; alcohol; beta-NAD; co-enzyme;  
 KM oxidize; optical isomer; (R)-2-octanol; 4-haloacetate ester;  
 KM (S)-4-halo-3-hydroxybutyrate ester; HMG-CoA reductase inhibitor;  
 KM D-carnitine; (R)-propoxybenzene.

OS Pichia finlandica.

PN W0200161014-A1.

PD 23-AUG-2001.

PF 15-FEB-2001; 2001WO-JP01082.

PR 16-FEB-2000; 2000JP-0043506.

PR 08-DEC-2000; 2000JP-0374593.

PA (DAIL ) DAICEL CHEM IND LTD.

PI Kudoh M, Yamamoto H;

DR WPI: 2001-596701/67.

DR N-PSDB; AAH43472.

PT (R)-2-octanol dehydrogenase for producing intermediates for HMG-CoA

PT reductase inhibitors or D-carnitine  
 PS Claim 9; Page 83-85; 97pp; Japanese.

CC This sequence represents (R)-2-octanol dehydrogenase which has the  
 CC following characteristics:  
 CC (1) produces ketones by oxidizing alcohol using beta-NAD as a  
 CC co-enzyme; and produces alcohols by reducing ketones using beta-NADH  
 CC as a co-enzyme; and  
 CC (2) preferentially oxidizes the optical isomer (R)-2-octanol; and  
 CC reduces 4-haloacetate ester to produce (S)-4-halo-3-  
 CC hydroxybutyrate ester.  
 CC (S)-4-halo-3-hydroxybutyrate esters are useful as intermediates  
 CC for HMG-CoA reductase inhibitors or D-carnitine. (R)-propoxybenzene  
 CC derivatives are particularly useful as intermediates for optical  
 CC isomers of ofloxacin (S)-(-)-9-fluoro-3-methyl-10-(4-methyl-1-  
 CC piperazinyl)-7-oxo-2,3-dihydro-7H-pyrido(1,2,3-de)(1,4) benzoxadine-  
 CC 6-carboxylic acid).

CC Sequence 254 AA:

Query Match 34.9%; Score 450; DB 22; Length 254;  
 Best Local Similarity 39.0%; Pred. No. 1.9e-35;  
 Matches 101; Conservative 46; Mismatches 100; Indels 12; Gaps 5;

OY 1 MSNRIDGKVAITITGTLGIGLAATKFEVBEKAVMITD-----RHSVGEKAKSVGTP 54  
 DB 1 MSYFNANKVAVTALSLGIGLAATKFEVBEKAVMITD-----RHSVGEKAKSVGTP 59  
 OY 55 DOIQFQHDSDSDGWTKLPDATEKAFGPVSTLVNNAIGAVNKSVEETTAEKRLAVN 114  
 DB 60 DNLHYVQADSSKEEDNKKLISSETLSAFGLDYVCANAGIATFTQTTDISTVWRKYTSIN 119  
 OY 115 LDGVFFGRGLGIQKRNKNGLGASIIINSSIEGFVGDPSLGAVNASKAVRIMSKSALDC 174  
 DB 120 LDGVFMDRLKLAQVFLSKNKPGLAIVMGSISHSYVAFGLSHCAAGKGLKLTQTALEY 179  
 OY 175 ALKDYDVARNVHPGYIKTPLVDDLPGAEEAMSOR-TKTPMGHIGEPNDIATICYTLASN 233  
 DB 180 A--SHGIRVANSVNPGYIKTPLVDDICP--KERLDKLVSLHPGRGKPEEVADAVAFICSD 235  
 OY 234 ESKFATGSEFVVDGTYAQ 252  
 DB 236 EATFINGISLVPDGGTYAQ 254

RESULT 6

AA556815 standard; Protein: 258 AA.

AC AA556815;

DT 31-MAR-2000 (first entry)

DE Bacillus D-arabinitol dehydrogenase.

KM D-arabinitol dehydrogenase; clinical diagnosis; mycosis.

OS Bacillus sp.

PN JP11332569-A.

PD 07-DEC-1999.

PF 26-MAY-1998; 98JP-0143637.

PR 26-MAY-1998; 98JP-0143637.

PA (IKED-) IKEDA SHOKKEN KK.

PA (NIPK ) NIPPON KAYAKU KK.

DR WPI: 2000-091353/08.

DR N-PSDB; AA246762, AA246763.

PT Novel Isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections -

PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, zyskind JW, Wall D, Trawick JD, Carr GJ

PD 12-OCT-2001.  
 PF 11-APR-2000; 2000FR-0004630.  
 XX 11-APR-2000; 2000FR-0004630.  
 PR  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA  
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
 PI  
 XX WPI; 2002-043418/06.  
 DR  
 XX  
 PT New nucleotide sequence useful in the identification of *Lactococcus*  
 PS *lactis* and related species -  
 XX  
 PS  
 XX Claim 6: SEQ ID NO 916; 2504pp; French.  
 CC The present invention is related to a *Lactococcus lactis* nucleotide  
 CC sequence (AB90521) and related proteins (AB85330-AB855621). The  
 CC nucleic acid sequence is useful in the detection and/or amplification of  
 CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or  
 CC related species. The proteins of the invention are useful for the  
 CC biosynthesis or biodegradation of a composition of interest. The  
 CC invention helps research in lactic bacteria, particularly useful in the  
 CC production of yogurt and cheese.  
 CC Note: The sequence data for this patent is based on equivalent patent  
 CC WO2001/7734 (published 18-OCT-2001) which is available in electronic  
 CC format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
 XX  
 XX Sequence 253 AA;  
 XX

Query Match	28.5%	Score 368;	DB 23;	Length 253;
Best Local Similarity	34.3%	Pred. No. 1.8e-27;		
Matches	86;	Conservative	48;	Mismatches 103;
			Indels	14;
			Gaps	4

[illegible]

RESULT 10	
AA27757	
ID	AA27757 standard; protein; 261 AA.
XX	
AC	AA27757;
XX	
DT	11-MAR-1993 (first entry)
XX	
DE	Glucose dehydrogenase.
XX	
GDH; mutant; recombinant; mass production; tetramer; thermostable	
XX	
OS	Bacillus megaterium.
XX	
XX	
Key	Location/Qualifiers
FT	Misc-difference 96
FT	/note- "site of Glu->Val or Lys mutation"

PT	18-JUN-2002 (first entry)	
XX		
DE	Lactobacillus rhamnosus acetoin dehydrogenase butA.	
XX		
KW	Enzyme; flavour; aroma; texture; nutritional; dairy manufacture; therapy;	
KW	fermentation process; anti-infection; rotavirus infection; heart disease;	
KW	infantile diarrhoea; lactose digestion; anti-cancer; autoimmune disorder;	
KW	anti-mutagenesis; immune system modulation; allergy; Helicobacter pylori;	
KW	antihypertensive effect; urogenital infection; hepatic encephalopathy;	
KW	bowel syndrome; endocarditis; transgenic microbe; acetoin dehydrogenase;	
KW	butA; EC 1.1.1.5.	
XX		
OS	Lactobacillus rhamnosus HN001.	
XX		
PN	WO200212506-A1.	
XX		
PD	14-FEB-2002.	
XX		
PF	08-AUG-2001; 2001WO-NZ00160.	
XX		
PR	08-AUG-2000; 2000US-0634238.	
XX	28-NOV-2000; 2000US-0724623.	
PA	(GENE-) GENESIS RES & DEV CORP LTD.	
PA	(VIAL-) VIALACTIA BIOSCIENCE NZ LTD.	
XX		
PI	Glenn M, Hawukala IJ, Bloksberg LN, Lubbers MW, Dekker J;	
PI	Christensson AC, Holland R, O'Coole PW, Reid JR, Coolbear T;	
DR	WPI: 2002-241760/29.	
PT	N-PSDB: AAD31886.	
PT		
PT	New polynucleotides and polypeptides from lactobacillus rhamnosus,	
PT	useful in e.g. improving the flavor, aroma, texture and health-related	
PT	benefits of milk-derived products, or in increasing properties of	
PT	microbes -	
XX		
XX	Claim 11; Fig 50; 257pp: English.	
XX		
CC	The present invention relates to a new isolated polynucleotide comprising	
CC	a sequence present in Lactobacillus rhamnosus strain HN001 and encoding a	
CC	polypeptide capable of modifying the flavour, aroma, texture, nutritional	
CC	and health benefits of milk-derived products, and/or survivability of	
CC	microbes in dairy manufacturing processes. The polynucleotides are useful	
CC	for improving the properties of microbes used in the manufacture of milk-	
CC	derived products such as cheeses, yogurt, fermented milk products, sour	
CC	milks and buttermilk; in modifying the flavour, aroma, texture and health	
CC	-related benefits of milk-derived products and in increasing the survival	
CC	of microbes during industrial fermentation processes. The bacteria may be	
CC	used to increase resistance to enteric pathogens and anti-infection	
CC	activity, including treatment of rotavirus infection and infantile	
CC	diarrhoea; aid in lactose digestion; as anti-cancer and anti-mutagensis;	
CC	liver cancer reduction; reduction of small bowel bacterial overgrowth;	
CC	immune system modulation and treatment of autoimmune disorders and	
CC	allergies; treatment of allergic responses to foods; reduction of blood	
CC	lipids and prevention of heart disease; antihypertensive effect;	
CC	prevention and treatment of urogenital infections, Helicobacter pylori,	
CC	or hepatic encephalopathy; treatment of inflammatory bowel disorder and	
CC	irritable bowel syndrome; modulation of endocarditis; and for improved	
CC	protein and carbohydrate utilization and conversion. The transgenic	
CC	microbial population can be administered to a mammal as an anti-	
CC	carcinogenic agent. The present sequence is Lactobacillus rhamnosus	
CC	acetoin dehydrogenase butA. The EC number for acetoin dehydrogenase is	
CC	1.1.1.5.	
XX		
XX		
XX	Sequence 261 AA;	
XX		
Query Match	28.4%;	Score 366.5; DB 23; Length 261;
Best Local Similarity	33.5%;	Pred. No. 2.6e-27;
Matches 85;	Conservative 50;	Mismatches 112; Indels 7; Gaps 4;

Db 1 MYRDLNGKAVAVVTGSGKIGACIAERFGDEBHAUVINYIADGHEGARKTADTVIKNGQAV 60  
 QY 59 FFQHDSDSDGKTKLEDATEKAFGPVSTLVNNGIAVNVKSVETTTAEEMKLLAVNDGV 118  
 Db 61 SIHADVSTGAEIGIASIVKTAESEFGRDVMVNNAGMEIKAPTHEVSLDDMKVIAINOTGV 120  
 QY 119 FFGTRLGIGQRMKNGKLGASIIIMSSIEGVPDPSLGAYNASKGAVRIMSKSALDCAKD 178  
 Db 121 FLGARAAALNYFLDHHQPGKITINISSVHEQIPMPTEFASVAAAKSVKLFETETIMEYV--N 178  
 QY 179 YDVRVNTVHPGYIKTPLYVDLPGAEAMSQRT--KTPMGHIGEPNDIAYICVYLAESK 236  
 Db 179 RGIRVNAIGPAGIETP--INAEKFAADKAOYDQIVAMIPQGRIGKPEVVAAGANWLASTE 237  
 QY 237 FATGSEFVVDGCT 250  
 Db 238 YVTGTTLEFDGCMT 251

## RESULT 12

AAR04044  
 ID AAR04044 standard; protein: 261 AA.

AC AAR04044;

DT 02-MAR-1993 (first entry)

DE Glucose dehydrogenase.

KW GDH; vector: E. coli; enzyme; assay; food.

XX Bacillus megaterium.

XX Key Location/Qualifiers

FT Misc-difference 22 /label= SER, ALA

FT Misc-difference 43 /label= ASP, GLU

FT Misc-difference 79 /label= ALA, SER

FT Misc-difference 95 /label= LEU, MET

XX JP02072878-A.

XX PD 13-MAR-1990.

XX PF 06-SEP-1988; 88JP-0223343.

XX PR 06-SEP-1988; 88JP-0223343.

XX (AMANO) AMANO PHARM KK.

XX WPI; 1990-121044/16.

XX PT Glucose dehydrogenase prodn. from Bacillus megaterium - by

XX PT culturing recombinant DNA-contg. transformants in nutrient

XX PT culture medium, used in food industry

XX PS Claim 1; 12pp; Japanese.

XX CC DNA encoding glucose dehydrogenase (GDH) derived from B. megaterium

XX CC encodes the amino acid sequence below. The DNA may be integrated

XX CC into a vector for replication in E. coli. A large amt. of GDH may

XX CC be produced at low cost. The GDH is used in clinical laboratory

XX CC tests in the food industry in enzymic assays for glucose

XX CC determination.

XX SQ Sequence 261 AA;

Query Match 28.3%; Score 364.5; DB 11; Length 261;

Best Local Similarity 33.3%; Pred. No. 4, 1e-27;

Matches 85; Conservative 49; Mismatches 112; Indels 9; Gaps 4;

QY 1 MSNRDLGKVAITTGCTLGIGLAIATKFEVGAQWMTDHRSDVGERAA----KSVGTPDQ 56  
 Db 1 MYRDLNGKAVAVVTGSGKIGACIAERFGDEBHAUVINYIADGHEGARKTADTVIKNGQAV 58  
 QY 57 IOFFQHDSDSDGKTKLEDATEKAFGPVSTLVNNGIAVNVKSVETTTAEEMKLLAVND 116  
 Db 59 AIAVAGDVTVESDVINLVQSAIKERKGLDVMINMGXENPVSSHENSLSDMKVVITDNL 118  
 QY 117 GFVFGTRLGIGQRMKNGKLGASIIIMSSIEGVPDPSLGAYNASKGAVRIMSKSALDCA 176  
 Db 119 GAFGSRRAIRKYFVENDIKGTVINMSVHEKIPMPLEFVHYAASKGKMLTETLAEVAP 178  
 QY 177 KDVVRVNTVHPGYIKTPLYVDLPGAEAMSQRTKTPMGHIGEPNDIAYICVYLAESK 235  
 Db 179 K--GIRVANNIGPAGIINTINAEKFAADKAOYDQIVAMIPQGRIGKPEVVAAGANWL 236  
 QY 236 KEATGSEFVVDGCT 250  
 Db 237 SYVTGTTLEFDGCMT 251

## RESULT 13

AAR24018  
 ID AAR24018 standard; protein: 261 AA.

AC AAR24018;

DT 27-NOV-1992 (first entry)

DE Thermostable glucose dehydrogenase from Bacillus megaterium.

XX Thermostable; increased heat stability; GDH.

XX Bacillus megaterium.

XX Key Location/Qualifiers

FT Misc-difference 22 /label= Ser, Ala

FT Misc-difference 43 /label= ASP, GLU

FT Misc-difference 79 /label= ALA, Ser

FT Misc-difference 95 /label= LEU, Met

FT Misc-difference 96 /note- "together with residues 252,253 may be

FT Misc-difference 252 /note- replaced by a thermostability-conferring

FT Misc-difference 253 /note- amino acid provided that taken together

FT Misc-difference 253 /note- they are not Glu, Gln, and Tyr respectively"

FT Misc-difference 253 /note- "together with residues 96,253 may be

FT Misc-difference 253 /note- replaced by a thermostability-conferring

FT Misc-difference 253 /note- amino acid provided that taken together

FT Misc-difference 253 /note- they are not Glu, Gln, and Tyr respectively"

XX US5114853-A.

XX 19-MAY-1992.

XX PF 22-SEP-1989; 89US-0410844.

XX PR 22-SEP-1988; 88JP-0237699.

XX (AMANO) AMANO PHARM KK.

XX Makino Y, Negoro S, Okada H, Urabe I;



DR WPI: 1992-192176/23.

XX DNA contg. glucose dehydrogenase gene of *Bacillus megaterium*

PT used to express heat stable protein in *Escherichia coli*, e.g. with

PT glutamine 96 replaced by alanine

XX

PS Claim 1; Column 13; 10pp; English.

XX

CC A DNA library was prepared in *E. coli* C600, and screened with probe

CC AA029704, resulting in 3 positive colonies. These were cultured, and

CC plasmid DNA prepared from them and cut with *EcoRI* and *SalI*. The

CC probe AA029704 was then used to probe a PAGE gel and found to

CC strongly hybridise with a 3.6kb fragment. The 3 colonies were all

CC thus found to have the same plasmid, and this was named PGDA1

CC was cut with *EcoRI* and *Sau3AI*, and a 930bp fragment produced and

CC sequenced, and was found to encode the sequence given. PGDA1 was

CC then cut with *EcoRI* and *PvuII* and a 1.5kb fragment isolated. This

CC was then cloned into expression vector pK223-3 and transformed into

CC *E. coli* JM105.

XX

SQ Sequence 261 AA;

Query Match 28.3%; Score 364.5; DB 13; Length 261;

Best Local Similarity 33.3%; Pred. No. 4.1e-27;

Matches 85; Conservative 49; Mismatches 112; Indels 9; Gaps 4;

QY 1 MSNRDGLGVAITGGTGLGIAIAIKFEVEGAKVMTDRHSDVGEKAA---KSVGTPDQ 56

DB 1 MYKDLKGVVVTSSSTGLGKXMAIRFATEKAKVNVYRKSXEANSVLEIEIKKVG--GE 58

QY 57 IQFQHDSDDEGWTKEFDETEKAFGPYSTLVNNAAGVANKSVETTTAEKRLAVALVD 116

DB 59 AIAVKGAVTVESDVIVLQSKIKFEKGLDVMINNAKXNPVSSHMSLSDNKKVYIDTMLT 118

QY 117 GVEFGTRGLGIQRMNKGASIIINNSIEGFVGPDLGAVNASKAVRIMSKAALDQAL 176

DB 119 GAFEGSRKALIKFEVENIKGTVINNSVHEKIPPLFHYAASKGKMLMETLALETAP 178

QY 177 KDVIVRNVTVHPGYIKTPL-VDDLPGAEANSSQRTKTPMGHIGEPNDIAYICVYLANS 235

DB 179 K-GIRVNNIGPAINTPINAEKFPADPOERADVESMIPMGYIGEPBEIAVAAMLASSEA 236

QY 236 KEATGSEFVVDGCT 250

DB 237 SYVTGITLFPADGGMT 251

RESULT 14

AAU34193

ID AAU34193 standard; Protein: 251 AA.

XX

AC AAU34193;

XX

DT 14-FEB-2002 (first entry)

XX

DE Staphylococcus aureus cellular proliferation protein #469.

XX

KW Antisense: prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN MO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001MO-US09180.

XX

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr CJ;

PI Yamamoto RT, Xu HH;

XX

DR WPI: 2001-611495/70.

XX

DR N-PSDB: AAS52052.

XX

PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX

PS Example 3; Seq ID No 5689; 511pp; English.

XX

The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*

CC pneumoniae, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 251 AA;

Query Match 28.2%; Score 364; DB 22; Length 251;

Best Local Similarity 35.5%; Pred. No. 4.3e-27;

Matches 91; Conservative 42; Mismatches 109; Indels 14; Gaps 6;

QY 5 LDGVAITGGTGLGIAIAIKFEVEGAKVMTDRHSDVGEKAAKSVGTPDQIFQDHS 64

DB 1 LENKVAVVTGASTGIGASAIALAQEGAYVLVDIAEAVSEYVOKIKSGKAKAYVVDI 60

QY 65 SDEGWTLPATEKAFGPYSTLVNNAAGI--AVNKSVEETTTAEKRLAVALNDGVFCT 122

DB 61 ASEQOINDFASEIRFOFHVVLFFNMGVDNAAGR-IHEYPTDVYDKIMNVDMGTFMT 119

QY 123 RLGIORMNKGGLGASIIINNSIEGFVGPDLGAVNASKAVRIMSKAALDQALDQYR 182

DB 120 KMLPLAMTK--GGSIVTSSFSGOADLYRSGINAAGAVINFTKSAIEIG-RD-GIR 175

QY 183 VNTVHPGYIKTPLVDLPGAE-----ANSORTKTPMGHIGEPNDIAYICVYLANS 235

DB 176 ANAIAFGTIERPLVDKLTGTSDEKGAFRBNQWMTPLGLGKPEEVGKLVFLASDES 235

QY 236 KEATGSEFVVDGCTA 251

DB 236 SFITGETIRIDGVA 251

RESULT 15

AAAB10740

ID AAAB10740 standard; Protein: 272 AA.

XX

AC AAAB10740;

XX

DT 26-JAN-2001 (first entry)

XX

DE *B. megaterium* glucose dehydrogenase GLGDH protein.

KM Glucose dehydrogenase, GlcDH; fusion protein; detection.

OS Bacillus megaterium.

PN MO200049039-A2.

PD 24-AUG-2000.

PF 08-FEB-2000; 2000MO-EP00978.

PR 19-FEB-1999; 99DE-1006920.

PA (MERE ) MERCK PATENT GMBH.

PI Linxweller W, Burger C, Poeschke O, Hofmann U, Wolf A.

DR MPI; 2000-558290/51.

DR N-PSDB; AAA97948.

PT Glucose dehydrogenase, fusion proteins, useful in expression systems for quick detection of foreign proteins by gel electrophoresis -

PS Disclosure; Page 54-55; 63pp; German.

CC This invention describes a novel recombinant fusion protein (A) comprising at least a first and a second amino acid sequence, characterized in that the first sequence has the biological activity of a glucose dehydrogenase. Glucose dehydrogenase is useful as a detector protein (in a detection system) for any type of recombinant protein/polypeptide in a fusion protein. Glucose dehydrogenase can be used to detect protein-protein interactions, where it corresponds to a partner of a recombinant protein/polypeptide. Expression vectors encoding the fusion proteins are useful to optimize the expression of the recombinant protein/polypeptide in a recombinant production method. The host cell is also useful for recombinant production of the protein/polypeptide. The fusion proteins containing glucose dehydrogenase can be quickly detected in SDS-PAGE gels. This sequence represents the Bacillus megaterium glucose dehydrogenase, GlcDH which is used in the construction of fusion proteins as described in the method of the invention.

CC Sequence 272 AA;

Query Match 28.28; Score 363.5; DB 21; Length 272;

Best Local Similarity 31.5%; Pred. No. 5.4e-27;

Matches 82; Conservative 51; Mismatches 108; Indels 19; Gaps 4;

```
QY 1 MSNRLDGKVAIIITGGTIGIGLAATKFEVBECAKMIT-----DRHSDVGEKAASV 51
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MYTDLMDKVVYITIGSGTGLRAMAVRFGQBEAKVYINTYNNBEELDAKKEVEAGGQAI 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 52 GTPDQIQFQGHSSDEDEGWTKLFDATEKAFGPSTLVNNAIGAVNKSVEETTTAEWRKL 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 -----YVGQDVTKREDVYVNLVQTAIKKEFTLWMINNAQENVPVSHSLDMNKKYI 113
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 112 AVNLDGVFEGTRIGIQAKKKGLGASTIINMSIGFVGDPSLGAVNASKGAVRINSKSA 171
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 114 DTNLTGAFGLSRAIKFYFENDIKGNVNNMSVHEMIPWPLFVHYASKGKMLMETLA 173
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 172 LDCALKDQVAVNTVHPGYIKTPL-VDDLPGAEEAMSGRTKTPMGHIGEPNDIATYCYL 230
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 174 LEYAPR--GIRVNNIGPGANNIPINAEKADPEGRADVESMIPMGYIOCKPREVAVAFL 231
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 231 ASNESKFAATGSEFVVDGTY 250
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 232 ASSQASVTGTITLFAADGWT 251
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: April 22, 2003, 17:15:27  
Job time : 39 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 17:14:51 ; Search time 14 Seconds

(Without alignments)  
529.613 Million cell updates/sec

Title: US-09-910-033A-2

Perfect score: 1290  
Sequence: 1 MSNRDGVKVAIIITGTTGIC.....NESKFRATSEPVVDGCTRAQ 252

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5  
Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1277	99.0	252	3	US-08-822-322-8 Sequence 8, Appl1
2	1277	99.0	252	4	US-09-466-109-8 Sequence 8, Appl1
3	1019	79.0	251	3	US-08-822-322-9 Sequence 9, Appl1
4	1019	79.0	251	4	US-09-466-109-9 Sequence 9, Appl1
5	407.5	31.6	255	4	US-08-815-225-4 Sequence 4, Appl1
6	395.5	30.7	256	4	US-09-504-358-14 Sequence 14, Appl1
7	395.5	29.1	256	4	US-09-134-001C-3562 Sequence 3562, Ap
8	376	29.1	257	4	US-09-134-001C-3562 Sequence 3562, Ap
9	358.5	27.8	261	4	US-09-468-738A-29 Sequence 29, Appl
10	358.5	27.8	261	4	US-09-940-019-29 Sequence 29, Appl
11	348.5	27.0	263	4	US-09-134-001C-4512 Sequence 4512, Ap
12	341	26.4	274	4	US-09-134-001C-4431 Sequence 4431, Ap
13	330.5	25.6	258	4	US-09-504-358-12 Sequence 12, Appl
14	330.5	25.6	258	4	US-09-954-314-12 Sequence 12, Appl
15	325	25.2	333	1	US-08-440-856A-4 Sequence 4, Appl1
16	324	25.1	337	1	US-08-440-856A-3 Sequence 3, Appl1
17	319.5	24.8	256	1	US-08-594-808B-7 Sequence 7, Appl1
18	313.5	24.3	290	4	US-09-134-001C-4339 Sequence 4339, Ap
19	312.5	24.2	313	4	US-09-413-814-9 Sequence 9, Appl1
20	308.5	22.9	257	4	US-09-287-097-2 Sequence 2, Appl1
21	307.5	23.8	267	4	US-09-134-001C-5042 Sequence 5042, Ap
22	305	23.6	273	6	5512669-4 Patent No. 5512669
23	302	23.4	262	4	US-09-363-189B-6 Patent No. 5512669
24	294	22.8	246	3	US-09-238-481-2 Sequence 2, Appl1
25	294	22.8	246	4	US-09-572-810A-2 Sequence 2, Appl1
26	293	22.7	246	6	5229279-7 Patent No. 5229279
27	291.5	22.6	283	4	US-09-367-012-1 Sequence 1, Appl1

28	291.5	22.6	283	4	US-09-777-157A-1 Sequence 1, Appl1
29	288	22.3	243	4	US-09-239-052-2 Sequence 2, Appl1
30	288	22.3	263	4	US-09-134-001C-3505 Sequence 3505, Ap
31	281	21.8	247	1	US-08-241-766-13 Sequence 13, Appl
32	279.5	21.7	315	3	US-08-793-035-9 Sequence 9, Appl1
33	279.5	21.7	315	3	US-08-793-035-10 Sequence 10, Appl
34	273	21.2	249	4	US-09-134-001C-4825 Sequence 4825, Ap
35	272	21.1	244	1	US-08-762-129-4 Sequence 4, Appl1
36	271	21.0	263	6	5229279-4 Patent No. 5229279
37	269	20.9	244	1	US-08-762-129-3 Sequence 3, Appl1
38	268	20.8	244	2	US-09-090-567-2 Sequence 2, Appl1
39	262.5	20.3	261	4	US-08-815-225-2 Sequence 2, Appl1
40	262.5	20.3	261	4	US-08-815-225-50 Sequence 50, Appl
41	262.5	20.3	261	4	US-09-347-878-50 Sequence 50, Appl
42	260	20.2	244	1	US-08-375-962B-13 Sequence 13, Appl
43	260	20.2	244	2	US-08-562-114B-13 Sequence 13, Appl
44	260	20.2	244	4	US-08-729-594A-13 Sequence 13, Appl
45	260	20.2	244	4	US-08-937-993-13 Sequence 13, Appl

## ALIGNMENTS

RESULT: 1  
US-08-822-322-8  
Sequence 8, Application US/08822322  
Patent No. 6037158  
GENERAL INFORMATION:  
APPLICANT: Hummel, Werner, Riebel, Betting  
TITLE OF INVENTION: Alcohol dehydrogenase and its use for  
TITLE OF INVENTION: enzymatic production of chiral hydroxy compounds  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822,322  
FILING DATE: 21-March-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 10 984  
FILING DATE: 21-March-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6037158man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: BOER 1076  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-9884  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-822-322-8  
Query Match 99.0% Score 1277, DB 3, Length 252;  
Best Local Similarity 99.2% Pred. No. 5.7e-129;  
Matches 250; Conservative 2; Indels 0; Gaps 0;  
OY 1 MSNRDGVKVAIIITGTTGICLAIAIKFVEKGVKIVITRHSVSGEAKKSVGPQIOFF 60  
DB 1 MSNRDGVKVAIIITGTTGICLAIAIKFVEKGVKIVITRHSVSGEAKKSVGPQIOFF 60

QY 61 QHDSDEGKTLFPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFF 120  
DB 61 QHDSDEGKTLFPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFF 120  
QY 121 GTRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 180  
DB 121 GTRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 180  
QY 181 VRVNTVHPGYIKTLPVDDLPGAEEMASORTKTPMGHIGEPNDIAYICVYLASNESKFATG 240  
DB 181 VRVNTVHPGYIKTLPVDDLPGAEEMASORTKTPMGHIGEPNDIAYICVYLASNESKFATG 240  
QY 241 SEFVVDGCGYTAQ 252  
DB 241 SEFVVDGCGYTAQ 252

RESULT 2  
US-09-466-109-8  
Sequence 8, Application US/09466109  
Patent No. 6225099  
GENERAL INFORMATION:  
APPLICANT: Hummel, Werner, Riebel, Bettina  
TITLE OF INVENTION: Alcohol dehydrogenase and its use for enzymatic production of chiral hydroxy compounds  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/466,109  
FILING DATE: 17-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/822,322  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6225099man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: BOER 1076  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-466-109-8

Query Match 99.0%; Score 1277; DB 4; Length 252;  
Best Local Similarity 99.2%; Pred. No. 5,7e-129;  
Matches 250; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSNRLDGKVAIITGTLGIGLAIAIKFVEBGAKVMTDRHSDVGEEKAKSVGTPDQIOFF 60  
DB 1 MSNRLDGKVAIITGTLGIGLAIAIKFVEBGAKVMTGRHSDVGEEKAKSVGTPDQIOFF 60  
QY 61 QHDSDEGKTLFPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFF 120  
DB 61 QHDSDEGKTLFPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFF 120

QY 121 GTRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 180  
DB 121 GTRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 180  
QY 181 VRVNTVHPGYIKTLPVDDLPGAEEMASORTKTPMGHIGEPNDIAYICVYLASNESKFATG 240  
DB 181 VRVNTVHPGYIKTLPVDDLPGAEEMASORTKTPMGHIGEPNDIAYICVYLASNESKFATG 240  
QY 241 SEFVVDGCGYTAQ 252  
DB 241 SEFVVDGCGYTAQ 252

RESULT 3  
US-08-822-322-9  
Sequence 9, Application US/08822322  
Patent No. 6037158  
GENERAL INFORMATION:  
APPLICANT: Hummel, Werner, Riebel, Bettina  
TITLE OF INVENTION: Alcohol dehydrogenase and its use for enzymatic production of chiral hydroxy compounds  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822,322  
FILING DATE: 21-March-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 10 984  
FILING DATE: 21-March-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6037158man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: BOER 1076  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-822-322-9

Query Match 79.0%; Score 1019; DB 3; Length 251;  
Best Local Similarity 80.5%; Pred. No. 2,8e-101;  
Matches 202; Conservative 14; Mismatches 35; Indels 0; Gaps 0;

QY 2 SNRLDGKVAIITGTLGIGLAIAIKFVEBGAKVMTDRHSDVGEEKAKSVGTPDQIOFFQ 61  
DB 1 TDRKKGVAIITGTLGIGLAIAIKFVEBGAKVITGRHADVGKXXXXXXXXXXXXXXX 60  
QY 62 HDSDDEGKTLFPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFFG 121  
DB 61 XXXXXXXXXXLDPATEKAFGPVSTLVNNAIGAVNKSVEETTTAEKRLLAVNLDCVFFG 120  
QY 122 TRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 181  
DB 122 TRGIGIORMKKNKGASIIINSSIEGFVGDPSIGAYNASKGAVRINSKSAALDCALKDYD 180



RESULT 6  
US-09-504-358-14  
; Sequence 14, Application US/09504358  
; Patent No. 6365376  
; GENERAL INFORMATION:  
; APPLICANT: Rouviere, Pierre E.  
; APPLICANT: Brzostowicz, Patricia C.  
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES  
; FILE REFERENCE: BC1001 US NA  
; CURRENT APPLICATION NUMBER: US/09/504,358  
; EARLIER APPLICATION NUMBER: 2000-02-15  
; EARLIER FILING DATE: 1999-February-19  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 14  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Brevibacterium sp HCU  
US-09-504-358-14

Query Match 30.7%; Score 395.5; DB 4; Length 256;  
Best Local Similarity 37.7%; Pred. No. 2.3e-34;  
Matches 98; Conservative 43; Mismatches 104; Indels 15; Gaps 6;  
QY 3 NRLDGKVAIITGGLTIGLAATKFEVGEAKVMTDRHSDVGEKAASVGTPOI-OFQ 61  
DB 2 NRLDGKVAIITGGLTIGLAATKFEVGEAKVMTDRHSDVGEKAASVGTPOI-OFQ 61  
QY 62 HDSSDEDMTKLFDATKAGPSTLVNNGI-AVNSVEETTAEMKRLAVNLGVFF 120  
DB 62 LDVSDSEVEIVSDIAKRFGALVNLNMGVGTADKPTHEIDERDLDLVSDVKGVFF 121  
QY 121 GTRLGIOBMKKGASIIINNSIEGFVDPDSLGAYNASKAVRIMSKAALCALDKYD 180  
DB 122 MTKHCIPYFKQAG-GGALVNFASIVGLVSGDELTPYHAKAVALTKODAV--TYGFSN 178  
QY 181 VRVNTVHPGYIKTPLY-----DLPAGEAMSORTKTPMGHIGEPNDIAYICVYLAS 232  
DB 179 IRVNAVAPGTLTLPVLVELSGRPGDGLGYTKLMG--AKHPLGRVGTPEEVAATLFLAS 236  
QY 233 NESKFATGSEFVVDGYYTAQ 252  
DB 237 EASFTTGAVLPVDGGYTAQ 256

RESULT 7  
US-09-954-314-14  
; Sequence 14, Application US/09954314  
; Patent No. 6465224  
; GENERAL INFORMATION:  
; APPLICANT: Rouviere, Pierre E.  
; APPLICANT: Brzostowicz, Patricia C.  
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES  
; FILE REFERENCE: BC1001 US NA  
; CURRENT APPLICATION NUMBER: US/09/954,314  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: 60/120,702  
; PRIOR FILING DATE: 1999-February-19  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 14  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Brevibacterium sp HCU  
US-09-954-314-14

Query Match 30.7%; Score 395.5; DB 4; Length 256;  
Best Local Similarity 37.7%; Pred. No. 2.3e-34;  
Matches 98; Conservative 43; Mismatches 104; Indels 15; Gaps 6;  
QY 3 NRLDGKVAIITGGLTIGLAATKFEVGEAKVMTDRHSDVGEKAASVGTPOI-OFQ 61  
DB 2 NRLDGKVAIITGGLTIGLAATKFEVGEAKVMTDRHSDVGEKAASVGTPOI-OFQ 61

DB 2 NRLDGKVAIITGGAAGKRIOSSELYASDAQVAVDVNEQGRATADAIRASGVANVWK 61  
QY 62 HDSSDEDMTKLFDATKAGPSTLVNNGI-AVNSVEETTAEMKRLAVNLGVFF 120  
DB 62 LDVSDSEVEIVSDIAKRFGALVNLNMGVGTADKPTHEIDERDLDLVSDVKGVFF 121  
QY 121 GTRLGIOBMKKGASIIINNSIEGFVDPDSLGAYNASKAVRIMSKAALCALDKYD 180  
DB 122 MTKHCIPYFKQAG-GGALVNFASIVGLVSGDELTPYHAKAVALTKODAV--TYGFSN 178  
QY 181 VRVNTVHPGYIKTPLY-----DLPAGEAMSORTKTPMGHIGEPNDIAYICVYLAS 232  
DB 179 IRVNAVAPGTLTLPVLVELSGRPGDGLGYTKLMG--AKHPLGRVGTPEEVAATLFLAS 236  
QY 233 NESKFATGSEFVVDGYYTAQ 252  
DB 237 EASFTTGAVLPVDGGYTAQ 256

RESULT 8  
US-09-134-001C-3562  
; Sequence 3562, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3562  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3562

Query Match 29.1%; Score 376; DB 4; Length 257;  
Best Local Similarity 33.7%; Pred. No. 2.8e-32;  
Matches 86; Conservative 59; Mismatches 94; Indels 16; Gaps 6;  
QY 4 RLDGKVAIITGGLTIGLAATKFEVGEAKVMTDRHSDVGEKAASVGTPOI-OFQ 61  
DB 11 KLTGKVAWVIGATGIGKMAEALQGANIVLADQSNIGOFATATITISQSGVKTSLK 70  
QY 62 HDSSDEDMTKLFDATKAGPSTLVNNGI-AVNSVEETTAEMKRLAVNLGVFF 121  
DB 71 LDITLHDEVNOIYDVVREKIDILVNNASISIODTEINISYEEMWKEINLSINGAFSV 130  
QY 122 TRLGIOBMKKGASIIINNSIEGFVDPDSLGAYNASKAVRIMSKAALCALDKYD 179  
DB 131 AQTGOMTEKSG-SMINVSVGLIANTQOOSSETSKAGVYTLTKSLAREMS--RY 187  
QY 180 DVAVNTVHPGYIKTPLYVLU;PGAEEAMSORTK--TPMGHIGEPNDIAYICVYLASNEK 236  
DB 188 GIKVNAIAPGYMRT-----IETEKILNDTEINTPMEVGEPEELAGITVYLASDASS 241  
QY 237 FATGSEFVVDGYYTA 251  
DB 242 FTQGSVEFNIDGYS 256

RESULT 9  
US-09-468-738A-29  
; Sequence 29, Application US/09468738A  
; Patent No. 6312953  
; GENERAL INFORMATION:  
; APPLICANT: Kimoto, No. 6312933jhlro  
; APPLICANT: Yamamoto, Hiroaki

Query Match	27.8%	Score 358.5	DB 4	Length 261
Best Local Similarity	33.9%	Pred. No. 2.2e-30		
Matches 87	Conservative 45	Mismatches 104	Indels 21	Gaps 5

RESULT 12  
US-09-134-001C-4431  
; Sequence 4431, Application US/09134001C  
; Patent No. 6380370

```

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4431
LENGTH: 274
TYPE: PRF
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4431

```

```

Query Match          26.48; Score 341; DB 4; Length 274;
Best Local Similarity 34.28; Pred. No. 1.8e-28;
Matches 88; Conservative 49; Mismatches 106; Indels 14; Gaps 7;

```

```

DB 4 RLDCRVALLTGTGIGLAIATKFEKAKVMTDRHSDVGEKAKSVG-TPDQIQFGH 62
DB 5 RLEMKIAVITGASGIGASAVALAEBAHVLAD-ISDQLETVQSTINDGCKTAAKV 63
DB 63 DSSDEGCTKLFDATEKAFGVPSTLVNNAIAVKNKSV-VEETTAEMRLLAVNLGDFEG 121
DB 64 DISDOKOVKPSKIEADFEFGHVDYFNNAGVNCAGRIHEXVEFEDIMAVDMGTFLV 123
DB 122 TRGIGORKNKNGIASIIMSSIEGFEVDPDLGAVNAKGAIRINSAALDCAIKDQV 181
DB 124 TKELPLPMKO--GGSIIITNASSGQADLYRSGYNAAKGVIINPKSIAIEYGRE--NI 179
DB 182 RVNTVHBYGTYKTPLVDDLPGA--EEA-----MSORTKTPMCHIGEPNIAIVCYVLASNE 234
DB 180 RANAIAAGTIEITPLVDNLAGTSDEAGOTFRFNCKWTPPLRLGTFPDEVGLVAFLASOD 239
DB 235 SKFATGSEFVVDGTYA 251
DB 240 SSFTTGTETIRIDGVA 256

```

```

RESULT 13
US-09-504-358-12
Sequence 12, Application US/09504358
Patent No. 6365376
GENERAL INFORMATION:
APPLICANT: Rouviere, Pierre E.
APPLICANT: Brzostowicz, Patricia C.
TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATE
FILE REFERENCE: BC1001 US NA
CURRENT APPLICATION NUMBER: US/09/504,358
CURRENT FILING DATE: 2000-02-15
EARLIER APPLICATION NUMBER: 60/120,702
EARLIER FILING DATE: 1999-February-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 258
TYPE: PRF
ORGANISM: Brevibacterium sp HCU
US-09-504-358-12

```

```

Query Match          25.68; Score 330.5; DB 4; Length 258;
Best Local Similarity 33.78; Pred. No. 2.2e-27;
Matches 91; Conservative 49; Mismatches 97; Indels 33; Gaps 8;

```

```

DB 1 MSNRLOGKVAIITGTLGIGLAIATKFEKAKVMTDRHSDVGEKAKSV--GTPDQIQ 58
DB 4 VYDSLGADVFLVYGAGGIGKATTTALAEGRVVLTDVEDAGSQAADVRRNTNGEIR 63
DB 59 FQHDSS-----DEGCTKLFDATEKAFGVPSTLVNNAIAVKNKSV-VEETTTAE 106

```

```

DB 64 FEPLDVTNPAVTECAOKLDDGM-----PYGLMANAGIAPSSAVDYSDEL 111
DB 107 WRKLLAVNLGVEFEFTRIGIORMKNKGASIIIMSSIEGF--VGDPSLCAVNAKGAVR 164
DB 112 WLRTVDINLVNGVFWCCREFGKRIARGRG--SVYTTSSIAGRTYSPERNAIYGATKAAYA 170
DB 165 IMKSAALDCAIKDQVRYNTVHBYGTYKTPLVDDLPGA--AEBAISORT--KTPMGHIEPND 222
DB 171 HLVLGLGVEMA--KTGVRNVAAPGYRTPTLEALKAESPETISEWTERIPNRLNDPSE 228
DB 223 IAYICVYLASNESKFAFGSEFVVDGTYAQ 252
DB 229 IADGVVFLMSNARGITGTVLHIDGTYAR 258

```

```

RESULT 14
US-09-954-314-12
Sequence 12, Application US/09954314
Patent No. 6465224
GENERAL INFORMATION:
APPLICANT: Rouviere, Pierre E.
APPLICANT: Brzostowicz, Patricia C.
TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIA
FILE REFERENCE: BC1001 US NA
CURRENT APPLICATION NUMBER: US/09/954,314
CURRENT FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/120,702
PRIOR FILING DATE: 1999-February-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 258
TYPE: PRF
ORGANISM: Brevibacterium sp HCU
US-09-954-314-12

```

```

Query Match          25.68; Score 330.5; DB 4; Length 258;
Best Local Similarity 33.78; Pred. No. 2.2e-27;
Matches 91; Conservative 49; Mismatches 97; Indels 33; Gaps 8;

```

```

DB 1 MSNRLOGKVAIITGTLGIGLAIATKFEKAKVMTDRHSDVGEKAKSV--GTPDQIQ 58
DB 4 VYDSLGADVFLVYGAGGIGKATTTALAEGRVVLTDVEDAGSQAADVRRNTNGEIR 63
DB 59 FQHDSS-----DEGCTKLFDATEKAFGVPSTLVNNAIAVKNKSV-VEETTTAE 106
DB 64 FEPLDVTNPAVTECAOKLDDGM-----PYGLMANAGIAPSSAVDYSDEL 111
DB 107 WRKLLAVNLGVEFEFTRIGIORMKNKGASIIIMSSIEGF--VGDPSLCAVNAKGAVR 164
DB 112 WLRTVDINLVNGVFWCCREFGKRIARGRG--SVYTTSSIAGRTYSPERNAIYGATKAAYA 170
DB 165 IMKSAALDCAIKDQVRYNTVHBYGTYKTPLVDDLPGA--AEBAISORT--KTPMGHIEPND 222
DB 171 HLVLGLGVEMA--KTGVRNVAAPGYRTPTLEALKAESPETISEWTERIPNRLNDPSE 228
DB 223 IAYICVYLASNESKFAFGSEFVVDGTYAQ 252
DB 229 IADGVVFLMSNARGITGTVLHIDGTYAR 258

```

```

RESULT 15
US-08-440-856A-4
Sequence 4, Application US/08440856A
Patent No. 5750873
GENERAL INFORMATION:
APPLICANT: DELLAPORTA, STEPHEN L.
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
FILE REFERENCE: BC1001 US NA
CURRENT APPLICATION NUMBER: US/08/440,856A
CURRENT FILING DATE: 2000-02-15
EARLIER APPLICATION NUMBER: 60/120,702
EARLIER FILING DATE: 1999-February-19
NUMBER OF SEQ ID NOS: 9
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER

```



STREET: 2000 PENNSYLVANIA AVE. N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,856A  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A.  
REGISTRATION NUMBER: 36,217  
REFERENCE/DOCKET NUMBER: 05463-20001.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1517  
TELEFAX: (202) 887-0763  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-440-856A-4

Query Match 25.2%; Score 325; DB 1; Length 333;  
Best Local Similarity 32.5%; Pred. NO. 1.3e-26;

Matches 92; Conservative 46; Mismatches 107; Indels 38; Gaps 7;

QY 1 MSNRIDGVAITGGTIGLAIATKFEVGEAKVMITDRHSDVGEKAASVGTPOIOFF 60  
DB 48 MPKRLGKVAIVTGAGRGGEIVRLFKHGAQVINDIDDAAGEALAAIG--PHYGV 105  
QY 61 QHDSDEDEGWTLEFDATEKAFGEVSTLVNNAI-----AVKSVETTTAEKRLAVNL 115  
DB 106 RCDVSEVEDVERAVERAARVARGRLDVCNNAAGVIGROTRAASILSPDAGEFDRVLRVNA 165  
QY 116 DGVFEGTIRIGIRMKNGKASIIINSSIEGFVGDPSLGAVNASKGAVRIMSKSALDCA 175  
DB 166 LGALGGMHAALAMTORRAG-SIIISVAVAGVLGGLCPHATASKHALVGLTKNA--CE 222  
QY 176 LKDYDVAVNTVHPGYIKTPLV-----DD-----LPGAEAA---MS 207  
DB 223 LGAHGIRVNCISPFVATPMLINAMRQHDASTADADADIDIDIAVPSDDEVKMEEV 282  
QY 208 QRTKTPMGIHIGEPNIAIYICVILASNESKFAATGSEFVYDGGYT 250  
DB 283 RGLATLKGAITLRPRDIAEAALPLASDSDRYISGHNLVVDGGVT 325

Search completed: April 22, 2003, 17:17:18  
Job time : 20 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 22, 2003, 17:16:32 ; Search time 25 Seconds  
(without alignments)  
807.708 Million cell updates/sec

Title: US-09-910-033A-2  
Perfect score: 1290  
Sequence: 1 MSNRLDGKVAITITGTTGIG.....NESKATGSEFVVDGCTAAQ 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications, AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1290	100.0	252	US-09-910-033A-2	Sequence 2, Appl
2	450	34.9	254	US-09-978-738-2	Sequence 2, Appl
3	419	33.5	253	US-09-931-186-15	Sequence 15, Appl
4	395.5	30.7	256	US-09-954-314-14	Sequence 14, Appl
5	369	28.6	272	US-09-815-242-12688	Sequence 12688, A
6	366.5	28.4	261	US-09-971-536-74	Sequence 74, Appl
7	364	28.2	251	US-09-815-242-5689	Sequence 5689, Ap
8	358.5	27.8	261	US-09-940-037A-29	Sequence 29, Appl
9	351	27.2	286	US-09-815-242-11860	Sequence 11860, A
10	345	26.7	259	US-09-931-186-17	Sequence 17, Appl
11	340	26.4	270	US-09-978-295A-468	Sequence 468, App
12	340	26.4	270	US-09-978-697-468	Sequence 468, App
13	340	26.4	270	US-09-978-192A-468	Sequence 468, App
14	340	26.4	270	US-09-999-832A-468	Sequence 468, App
15	340	26.4	270	US-09-978-189-468	Sequence 468, App
16	340	26.4	270	US-09-978-608A-468	Sequence 468, App
17	340	26.4	270	US-09-978-191A-468	Sequence 468, App
18	340	26.4	270	US-09-978-403A-468	Sequence 468, App
19	340	26.4	270	US-09-978-564A-468	Sequence 468, App

20	340	26.4	270	US-09-978-585A-468	Sequence 468, App
21	340	26.4	270	US-10-017-081A-468	Sequence 468, App
22	340	26.4	270	US-09-978-824-468	Sequence 468, App
23	340	26.4	270	US-09-981-915A-468	Sequence 468, App
24	340	26.4	270	US-09-999-833A-468	Sequence 468, App
25	340	26.4	270	US-10-167-749-468	Sequence 468, App
26	340	26.4	270	US-09-918-585A-468	Sequence 468, App
27	340	26.4	270	US-09-978-423A-468	Sequence 468, App
28	340	26.4	270	US-10-013-921A-468	Sequence 468, App
29	340	26.4	270	US-09-978-193A-468	Sequence 468, App
30	340	26.4	270	US-10-013-929A-468	Sequence 468, App
31	340	26.4	270	US-10-016-177A-468	Sequence 468, App
32	336.5	26.1	271	US-09-815-242-13387	Sequence 13387, A
33	336.5	26.1	272	US-09-815-242-13679	Sequence 13679, A
34	330.5	25.6	258	US-09-954-114-12	Sequence 12, Appl
35	324	25.1	274	US-09-976-059-17	Sequence 17, Appl
36	321.5	24.9	251	US-09-922-501-16	Sequence 16, Appl
37	317	24.6	258	US-10-156-055A-1	Sequence 1, Appl
38	311.5	24.1	253	US-09-815-242-11842	Sequence 11842, A
39	306	23.7	245	US-09-815-242-10921	Sequence 10921, A
40	305	23.6	262	US-09-815-242-10597	Sequence 10597, A
41	302	23.4	262	US-09-802-853-6	Sequence 6, Appl
42	301.5	23.4	276	US-09-944-160-24	Sequence 24, Appl
43	299.5	23.2	258	US-09-738-626-6436	Sequence 6436, Ap
44	299	23.2	242	US-09-815-242-10974	Sequence 10974, A
45	294	22.8	246	US-09-815-242-5461	Sequence 5461, Ap

## ALIGNMENTS

RESULT 1  
US-09-910-033A-2  
; Sequence 2, Application US/09910033A  
; Publication No. US20030054520A1  
GENERAL INFORMATION:  
APPLICANT: RIBBEL, Bettina  
APPLICANT: HUMMEL, Werner  
TITLE OF INVENTION: RECOMBINANT ENZYMES HAVING IMPROVED NAD(H) ACCEPTANCE  
FILE REFERENCE: 21021205  
CURRENT APPLICATION NUMBER: US/09/910.033A  
PRIOR FILING DATE: 2001-07-23  
PRIOR APPLICATION NUMBER: DE 10037101.9  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Lactobacillus brevis  
US-09-910-033A-2

Query Match: 100.0% Score 1290; DB 9; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1e-113;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSNRLDGKVAITITGTTGIGLAITKFEEGAKVITDRHSVDGEKAKSVGTPDIOFF 60  
1 MSNRLDGKVAITITGTTGIGLAITKFEEGAKVITDRHSVDGEKAKSVGTPDIOFF 60  
1 MSNRLDGKVAITITGTTGIGLAITKFEEGAKVITDRHSVDGEKAKSVGTPDIOFF 60  
OHSSDSDGWTKLPEDATEKAFGPSTLVNNGIAVNSVEETTTAEMRKLAVNLGVFF 120  
OHSSDSDGWTKLPEDATEKAFGPSTLVNNGIAVNSVEETTTAEMRKLAVNLGVFF 120  
GTRIGIORMKKKGAGASTINMSIEGFVGDPSLGAVNMSKGVRIKMSKAAIDCAKDVD 180  
GTRIGIORMKKKGAGASTINMSIEGFVGDPSLGAVNMSKGVRIKMSKAAIDCAKDVD 180  
VRNTVTPGKIKTPVVDLPAGEAMSORRTKPMKIGEPNDIAIYICVLASNESFATG 240  
VRNTVTPGKIKTPVVDLPAGEAMSORRTKPMKIGEPNDIAIYICVLASNESFATG 240  
VRNTVTPGKIKTPVVDLPAGEAMSORRTKPMKIGEPNDIAIYICVLASNESFATG 240

OY 241 SEFVVDGYYTAQ 252  
Db 241 SEFVVDGYYTAQ 252

## RESULT 2

US-09-978-758-2  
Sequence 2, Application US/09978758  
Publication No. US20020192783A1  
GENERAL INFORMATION:  
APPLICANT: Kudooh, Masatake  
APPLICANT: Yamamoto, Hiroaki  
TITLE OF INVENTION: (R)-2-OCTANOL DEHYDROGENASES, METHODS  
TITLE OF INVENTION: FOR PRODUCING THE ENZYMES, AND  
TITLE OF INVENTION: METHODS FOR PRODUCING ALCOHOLS USING THE ENZYMES  
FILE REFERENCE: 06501-090001  
CURRENT APPLICATION NUMBER: US/09/978,758  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: PCT/JP01/01082  
PRIOR FILING DATE: 2001-02-15  
PRIOR APPLICATION NUMBER: JP 2000-374593  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: JP 2000-43506  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Pichia finlandica  
US-09-978-758-2

Query Match 34.9%; Score 450; DB 9; Length 254;  
Best Local Similarity 39.0%; Pred. No. 1,1e-34;  
Matches 101; Conservative 46; Mismatches 100; Indels 12; Gaps 5;

OY 1 MSNRDGVAAITGGTIGLAIATKFEVGEAKVMTD-----RHSDVGEAKASVGP 54  
Db 1 MSYNHNKVAAYVTGALSGISLVAKKFLQLGAKYISDVSGEKKHETVVALKAQNLMT- 59  
OY 55 DOIQFOHDSDEDMGTKLFDATKAFGPVSTLVNNAICAVNKSVEETTTAKMKLLAVN 114  
Db 60 DNLHYVQADSSKEEDNKLLISFTLATFGGLDVCANAGICFAPTHETPFDVWKVIAVN 119  
OY 115 LDGYEFTGRLGIQRMKNGLASITMSSIEGFGVDPGLGAVNASKGAVRIMSKAALDC 174  
Db 120 LAGVFLDLKLAINWLEKSPGVIVNMGSHSVFAAPCLAHYGAAGVKLLITQTLALEY 179  
OY 175 ALKOYDVAVNVHGYITPLVDLPGAEEMASORTKT-PMCHIGEPPNDIAYICVYLASN 233  
Db 180 A-SHGIRVNVNPGYISTPLIDEVP--KERLDKLVSLHPIGRLGRPEVADAVAFLSQ 235  
OY 234 ESKFATGESEFVVDGYYTAQ 252  
Db 236 EATFINGVSLFVVDGYYTAQ 254

## RESULT 3

US-09-931-186-15  
Sequence 15, Application US/09931186  
Patent No. US20020132319A1  
GENERAL INFORMATION:  
APPLICANT: ABBRO, MELVYN A.  
APPLICANT: AGREE, CHARLES S.  
APPLICANT: AUST, ROBERT  
APPLICANT: KISSINGER, CHARLES R.  
APPLICANT: MARGOSIAK, STEVE  
APPLICANT: MENG, JERRY J.  
APPLICANT: PELLETIER, LAURA A.  
APPLICANT: REJTO, PAUL A.  
APPLICANT: SHOMALTER, RICHARD E.  
APPLICANT: TEMPCZYK-RUSSELL, ANNA  
APPLICANT: THOMSON, JIM

APPLICANT: VANDERPOOL, DARIN  
APPLICANT: VILLAFRANCA, J. ERNEST  
TITLE OF INVENTION: PEPTIDE MUTANT OF HUMAN ERAB/HADH2, ITS X-RAY CRYSTAL  
STRUCTURE, AND MATERIALS AND METHOD FOR IDENTIFICATION  
TITLE OF INVENTION: OF INHIBITORS THEREOF  
FILE REFERENCE: 0125-0049  
CURRENT APPLICATION NUMBER: US/09/931,186  
CURRENT FILING DATE: 2001-08-17  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 253  
TYPE: PRT  
ORGANISM: Steptomyces hydrogenans  
US-09-931-186-15

Query Match 32.5%; Score 419; DB 10; Length 253;  
Best Local Similarity 39.1%; Pred. No. 9,1e-32;  
Matches 97; Conservative 42; Mismatches 103; Indels 6; Gaps 4;

OY 3 NRIDGVAAITGGTIGLAIATKFEVGEAKVMTDRHSDVGEAKASVGPDOIQFQH 62  
Db 1 NDLGKTVIITGGARGLGAARQAVAGARVLAVDLDEGAATARELG--DAARYHL 58  
OY 63 DSDDEDMTKLFDATKAFGPVSTLVNNAICAVNKSVEETTTAKMKLLAVNLDGVFPGT 122  
Db 59 DVTIEEDMORVAVAREEGSVGLVNNAGISTGMPLETESVERFRKVEINLTGVFQGM 118  
OY 123 RLGIQRMKNGLASITMSSIEGFGVDPGLGAVNASKGAVRIMSKAALDCALDKYDR 182  
Db 119 KTVIPAMKQAG--CGSIVNISSAAGLGLALTSYSGASKGKGLSLAIVE--LGTDRIR 175  
OY 183 VNTVHPGYITPLVDLPGAEEMASORTKT-PMCHIGEPPNDIAYICVYLASNESKFGTSE 242  
Db 176 VNSVHPGMYTFPTMAE--TGIRQEGNYPTMPGRVCEPEIGAVYKLLSDTSYTTGAE 234  
OY 243 FVVDGYYT 250  
Db 235 LAVDGGWT 242

## RESULT 4

US-09-954-314-14  
Sequence 14, Application US/09954314  
Patent No. US20020127866A1  
GENERAL INFORMATION:  
APPLICANT: Brozowicz, Patricia C.  
APPLICANT: Rouviere, Pierre E.  
TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIA  
FILE REFERENCE: BC1001 US NA  
CURRENT APPLICATION NUMBER: US/09/954,314  
CURRENT FILING DATE: 2001-09-17  
PRIOR APPLICATION NUMBER: 60/120,702  
PRIOR FILING DATE: 1999-February-19  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 14  
LENGTH: 256  
TYPE: PRT  
ORGANISM: Brevibacterium sp HCU  
US-09-954-314-14

Query Match 30.7%; Score 395.5; DB 10; Length 256;  
Best Local Similarity 37.7%; Pred. No. 1,5e-29;  
Matches 98; Conservative 43; Mismatches 104; Indels 15; Gaps 6;

OY 3 NRIDGVAAITGGTIGLAIATKFEVGEAKVMTDRHSDVGEAKASVGPDOI-OFQF 61  
Db 2 NRIDGVAAITGGAGAGMGRIOSELVASEGAVAVDVNEOECRATADADIRASGVANYWK 61  
OY 62 HDSDEDMTKLFDATKAFGPVSTLVNNAIC-ANVKSVEETTTAKMKLLAVNLDGVF 120  
Db 62 LDVDESEVELVSDIARFGAIVNLYNAGYTGADKFTHEIDERDLDVLVSDVKGVEF 121

OY 121 CTRLCIQRKNNKGLGSIINMSIEGFVCPDPSLCAYNASKGAVRIMSKSALDCLKDYD 180  
Db 122 MTKHCIPYFKQAG-GGAIVFASIIYGLVSGOELPPYNAKGAVALFTQDQAV--TYGSPN 178  
OY 181 VAVNTVHPGICYITPLV-----DLPAGEAAMSORTKPTMGHIGCEPNDIAVICVYLAS 232  
Db 179 IAVNNAVAPCTILTPVLKELGSRCPDGLDGYTKLMG--AKHFLRGVGTPEEVAATLFLAS 236  
OY 233 NESKPTGSEFVVDGCTTAQ 252  
Db 237 EASFTTGAIVLPDGGCTTAQ 256

## RESULT 5

US-09-815-242-12688  
Sequence 12688, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011a  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12688  
LENGTH: 272  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12688

Query Match 28.6%; Score 369; DB 10; Length 272;  
Best Local Similarity 35.8%; Pred. No. 5,1e-27;  
Matches 92; Conservative 42; Mismatches 109; Indels 14; Gaps 6;

OY 4 RLDGKVAIITGGTGLGIAIATKFEVEGAKVMTDRHSVDGEKAKSVGTPODIOFQHD 63  
Db 3 RLENKVAAYVTGASTGIGQASAIALAGCAVYLVADIAEAVSETYDKIKNSGDKAKAYVD 62  
OY 64 SDEDCWTLKFDATKAFGVSTLVNNAIGT--AVNKSVEETTTAEMRKLAVNLGCVFFG 121  
Db 63 IASGOIDNFASIRFOCFHVDVLFNNAGVDNAGR--IHEPTVDYDKIMNVDMGTFLM 121  
OY 122 TRLCIQRKNNKGLGSIINMSIEGFVCPDPSLCAYNASKGAVRIMSKSALDCLKDYD 181  
Db 122 TMMLPLMNTK--GGSIVNTSSFSQADLXRSYMAKGAIVNTKSAIAIEYG-RD-GI 177  
OY 182 RVNTVHPGICYITPLVDDLPAGEE-----AMSORTKPTMGHIGCEPNDIAVICVYLASNE 234  
Db 178 RANAIAPGTIEPTPLVVKLTGTSEDKKAPRENQKMMPTLCRLGKPEVGLVFLASDE 237

OY 235 SKPATGSEFVVDGCTTA 251  
Db 238 SSFTGTETIRIDGCVMA 254

## RESULT 6

US-09-971-536-74  
Sequence 74, Application US/09971536  
Patent No. US20020159976A1  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Havukkala, Ilkka  
APPLICANT: Bloksberg, Leonard  
APPLICANT: Lubbers, Mark  
APPLICANT: Dekker, James  
APPLICANT: Christenson, Anna  
APPLICANT: Holland, Ross  
APPLICANT: O'Toole, Paul  
APPLICANT: Reid, Julian  
TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Me  
FILE REFERENCE: 1043c2  
CURRENT APPLICATION NUMBER: US/09/971,536  
PRIOR FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238  
PRIOR FILING DATE: 2000-08-08  
PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: PCT/NZ01/00160  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 74  
LENGTH: 261  
TYPE: PRT  
ORGANISM: Lactobacillus rhamnosus  
US-09-971-536-74

Query Match 28.4%; Score 366.5; DB 9; Length 261;  
Best Local Similarity 33.5%; Pred. No. 8.2e-27;  
Matches 85; Conservative 50; Mismatches 112; Indels 7; Gaps 4;

OY 1 MSNRLDKVAIITGGTGLGIAIATKFEVEGAKVMT--DRHSVDGEKAKSVGTPODIO 58  
Db 1 MYRDLNCKVAAYVTGSGSKIGAGIAERFGEHMAVINYLGDEGAKAKTADTVIKNGQAV 60  
OY 59 FFOHDSDEDCWTLKFDATKAFGVSTLVNNAIGTAVNKSVEETTTAEMRKLAVNLGCV 118  
Db 61 SIHADVSTEGIASLVKTAESEFGRLDVWVNNAGMEIKAPTHEVSLDDNNKVIAINQTV 120  
OY 119 FFCRLCIGRNNKGLGSIINMSIEGFVCPDPSLCAYNASKGAVRIMSKSALDCLKDYD 178  
Db 121 FLGRALALNTPFLDHPONINIISVHQDIPPTASVAAAGSVLFTFETIMETA--N 178  
OY 179 YDVAVNTVHPGICYITPLVDDLPAGEAAMSORT--KTPMGHIGCEPNDIAVICVYLASNE 236  
Db 179 RGIHVNAIGPAIETP--INAEKADKAOYDQTVAMIPOGRICKPEVDAAGAAVLASTESS 237  
OY 237 FATGSEFVVDGCTT 250  
Db 238 YVTGTTLFIDGCT 251

## RESULT 7

US-09-815-242-5689  
Sequence 5689, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.

```

1 RESULT 9
2 US-09-815-242-11860
3 : Sequence 11860 Application US/09815242
4 : Patent No. US20020061569A1
5 :
6 GENERAL INFORMATION:
7 APPLICANT: Haselbeck, Robert
8 APPLICANT: Ohlsen, Karl L.
9 APPLICANT: Zyskind, Judith W.
10 APPLICANT: Wall, Daniel
11 APPLICANT: Trawick, John D.
12 APPLICANT: Carr, Grant J.
13 APPLICANT: Yamamoto, Robert T.
14 APPLICANT: Xu, H. Howard
15 TITLE OF INVENTION: Identification of Essential Genes In
16 TITLE OF INVENTION: Prokaryotes
17 FILE REFERENCE: ELITRA.011A
18 CURRENT APPLICATION NUMBER: US/09/815,242
19 CURRENT FILING DATE: 2001-03-21
20 PRIOR APPLICATION NUMBER: 60/191,078
21 PRIOR FILING DATE: 2000-03-21
22 PRIOR APPLICATION NUMBER: 60/206,848
23 PRIOR FILING DATE: 2000-05-23
24 PRIOR APPLICATION NUMBER: 60/207,727
25 PRIOR FILING DATE: 2000-05-26
26 PRIOR APPLICATION NUMBER: 60/242,578
27 PRIOR FILING DATE: 2000-10-23
28 PRIOR APPLICATION NUMBER: 60/253,625
29 PRIOR FILING DATE: 2000-11-27
30 PRIOR APPLICATION NUMBER: 60/257,931
31 PRIOR FILING DATE: 2000-12-22
32 PRIOR APPLICATION NUMBER: 60/269,308
33 PRIOR FILING DATE: 2001-02-16
34 NUMBER OF SEQ ID NOS: 14110
35 SOFTWARE: FastSeq for Windows Version 4.0
36 SEQ ID NO 11860
37 LENGTH: 286
38 TYPE: prt
39 ORGANISM: Pseudomonas aeruginosa

```



PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 26.4%; Score 340; DB 9; Length 270;  
Best Local Similarity 37.1%; Pred. No. 2.7e-24;  
Matches 95; Conservative 44; Mismatches 93; Indels 24; Gaps 10;

4 RUDGVAIIITGTLGIALINATKFEKGAKWITDRHSOVGEKAASVGTPOIOIOPFOHD 63

Db 6 RYAGKVVVYTGGRGIGAGIVRAFNVCARVVICDKDESGRALEQEL--PGAN-FILCD 62  
Qy 64 SDEDCNFKLPDATEKAFGVSTLVNAG-IAVNVSEETTAAEMRKLAVNLDGVEFGT 122  
Db 63 VTQEDDKVLTVSEIRFGHLDVCVNNAGHPPQRPPEETSAQGFROLLEINTLGTVLT 122  
Qy 123 RLQIQM-KKKGAGSIINNSIEGFYGDPSLGYNNSKAVRIMSSAALDCLAKQYD 181  
Db 123 KLLAPYLRISQ--NVINISSLVGAIGQAVPYVATKGAVTMTALALD--ESPFGV 177  
Qy 182 RVNTHRGYIKTPLVDDL-----PGA---EAMSORTKTPMGHIGEPNDIAYICVYLAS 232  
Db 178 RVNCISPGNITWTFMLMEBLALMPDPRATIREGMLAQ-----PLGIMGQPAEYGAANVFLAS 233  
Qy 233 NESKATGSEFVYDGC 248  
Db 234 -EAMFCTGIELLVTCG 248

## RESULT 12

US-09-978-697-468  
Sequence 468, Application US/09978697  
Patent No. US20020169284A1

## GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Guiney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James:  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC27  
CURRENT APPLICATION NUMBER: US/09/978,697  
PRIOR FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569



PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083366  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
Query Match 26.4%; Score 340; DB 9; Length 270;  
Best Local Similarity 37.1%; Pred. No. 2.7e-24;  
Matches 95; Conservative 44; Mismatches 93; Indels 24; Gaps 10;  
QY 4 RLDDKVAIIITGTLGIGIAITKFEEGAKWITRHSDDVGEKAAKSVGTPOQIFQND 63  
DB 6 RYAKVVVVTGGGIGIGIYIRAFVNSGARVYICKDESGRALLDEL--PGAV-FILCD 62  
QY 64 SSDEDCWTKLFDATEKAFGPVSTLVNNG-IAVNSVEETTAERKLLAVNLDDVFECT 122  
DB 63 VTQEDDVKTLVSETIRREGRDCVNNAGHPPQRPPEETSAGQFROLLELNLGTYTLT 122  
QY 123 RLGIQRM-KNKGIGASIIINMSIGFVGDPSLGAVNASKGAVRIMSAAALDALKDYD 181  
DB 123 KLALPYLRKSOG--NVINISLVGAIGQAQAVPVATKGAVTMTKALALD--ESPVG 177  
QY 182 RVNTVHPGYIKTPLYDDJ-----PGA---EAMSQRTKTPMGHIGEPNDIAICVLYAS 232  
DB 178 RVNCISPGNIWTFPLMEELALMPDPRAIIRGMLAQ-----FLGRMGQPAEVGAANVFLAS 233  
QY 233 NESKFKTSGSEFVVDG 248  
DB 234 -EANFCTGIELVYTG 248  
RESULT 13  
US-09-978-192A-468  
Sequence 468, Application US/09978192A  
Patent No. US2002017753A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kiljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC9  
CURRENT APPLICATION NUMBER: US/09/978, 192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311

[illegible]

1 CURRENT APPLICATION NUMBER: US/09/999,832A  
2  
3 CURRENT FILING DATE: 2001-10-24  
4  
5 PRIOR APPLICATION NUMBER: 09/915855  
6  
7 PRIOR FILING DATE: 2001-07-30  
8  
9 PRIOR APPLICATION NUMBER: 60/062250  
10  
11 PRIOR FILING DATE: 1997-10-17  
12  
13 PRIOR APPLICATION NUMBER: 60/064249  
14  
15 PRIOR FILING DATE: 1997-11-03  
16  
17 PRIOR APPLICATION NUMBER: 60/065311  
18  
19 PRIOR FILING DATE: 1997-11-13  
20  
21 PRIOR APPLICATION NUMBER: 60/066364

PRIOR APPLICATION NUMBER: 60/077450  
 PRIOR FILING DATE: 1998-03-10  
 PRIOR APPLICATION NUMBER: 60/077632

;; PRIOR APPLICATION NUMBER: 60/077641  
;; FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004

;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936

;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939

;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656

;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689

;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728

PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920

PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105

PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165

PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327

PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070

PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/091071

PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195

1	PRIOR APPLICATION NUMBER: 60/085333
2	PRIOR FILING DATE: 1998-05-13
3	PRIOR APPLICATION NUMBER: 60/085338
4	PRIOR FILING DATE: 1998-05-13
5	PRIOR APPLICATION NUMBER: 60/085342
6	PRIOR FILING DATE: 1998-05-13
7	PRIOR APPLICATION NUMBER: 60/085358
8	PRIOR FILING DATE: 1998-05-13
9	PRIOR APPLICATION NUMBER: 60/085700
10	PRIOR FILING DATE: 1998-05-15
11	PRIOR APPLICATION NUMBER: 60/085688
12	PRIOR FILING DATE: 1998-05-15
13	PRIOR APPLICATION NUMBER: 60/085579
14	PRIOR FILING DATE: 1998-05-15
15	PRIOR APPLICATION NUMBER: 60/085580
16	PRIOR FILING DATE: 1998-05-15
17	PRIOR APPLICATION NUMBER: 60/085577
18	PRIOR FILING DATE: 1998-05-15
19	PRIOR APPLICATION NUMBER: 60/085704
20	PRIOR FILING DATE: 1998-05-15
21	PRIOR APPLICATION NUMBER: 60/085697
22	PRIOR FILING DATE: 1998-05-15
23	PRIOR APPLICATION NUMBER: 60/085697
24	PRIOR FILING DATE: 1998-05-15

```

OY      4 RLDDKVAIIITGGTGLGICLAIAITKVEEGKAWIMIDRHSIDVCEKAASVGPDDIOIQPHD 63
Db      6 RYAKKVVVVYVYGGGIGAGIYRAEYVNSGARVYICDDKDESGRALLEQEL--PGAV-FILCD 62
OY      64 SSDEDMKTLFDATEKAFPGVSTLVNNAG-TAIVNKSVYEETTAEMRKLAVNLIDGVEFGT 122
Db      63 VTQDDKVTALVSEYIRRFGRIDCVVNNAGHHHPQPREETSAGFROLLELTNLIGTYTLT 122
OY      123 RLGIQRW-KNKGGLASIIINNSIEGFAGDPBSLGAVYNSKSGAVRITMSASALDCLAKRYDV 18
Db      123 KLALEPYRKSSGG---ANINISLVLGAIGQAQAVPYATKGAIVMTATLALAD--ESPEGV 177
OY      182 RVNTVHHGQYIKTLPVLDD-----PGA-----EANSQRTKTPMGHIGEPNDIAYICYVLAS 233
Db      178 RVNCISPGNITWPLMEELALALPMPRRATIRGEMLAQ---PLGRMGQPAEYGAAGAAVFLAS 233
OY      233 NESKATGSEFVYDGC 248
Db      234 -EANFTGIELLVYTG 248

```

RESULT 15  
US-09-978-189-468  
Sequence 468, Application US/09978188  
Publication No. US20030004102A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertslen, Mary E.  
APPLICANT: Goddard, Audrey J.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gunney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavani, Ivar J.  
APPLICANT: Kno, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paonli, Nicholas F.

APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C7  
CURRENT APPLICATION NUMBER: US/09/978,189  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640

```
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
```

Query Match 26.48; Score 340; DB 9; Length 270;

Best Local Similarity 37.18; Pred. No. 2.7e-24;

```
Matches 95; Conservative 44; Mismatches 93; Indels 24; Gaps 10;

QY 4 RLDGKVAITGCTGIGIAITKFEYEGAKVMTDRHSDVGEKAASVGPDIQIOFFQHD 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 6 RYAKGVVVYTGCGIGAGIYRAFVNSGARVYICDKDESGRALOEEL--PGAV-FILCD 62
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 64 SSDEGWTKLFDATKAFGVPSTLVNMG-IAVNSVEETTAEMRKLAVNLGVFFGT 122
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 VTQEDDVTVLSETRFRGRDLCVNNAGHPPQRPETSAOGFROLLELNLGYTTLT 122
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 RLGIORM-KNKGCLASITNMSIEGFVDPISGATNASKGAVRIMSKSALDCAKDXYD 181
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 KLALPYLKRKSG---NVNMISSLVGAIGQAQVPPVATKGAVTAMTKALALD--ESP 177
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 RVNTVHPGYITPLVDDL-----PGA---EAMSORTKTPMGIHGEPNDAIYCYLIAS 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 178 RVNCTSPGNITPLMEELALMPDPRAITREGMLAQ-----PLGRMGQPAVEGAAVFLAS 233
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 NESKPAIGSEFVVDG 248
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 234 -EANFCTGIELLVYTG 248
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: April 22, 2003, 17:21:49  
Job time : 27 secs



R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S. Nature 193, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; MOID:98295987; PMID:9634230  
 A:Accession: H70758  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-260 <COL>  
 A:Cross-references: GB:Z74025; GB:AL123456; NID:93261586; PIDN:CAA98414.1; PID:93261591  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: fabB3  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 F:8-184/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match	36.0%	Score	464.5	DB 2	Length	260			
Best Local Similarity	42.2%	Pred. No.	1,2e-29						
Matches	106	Conservative	44	Mismatches	90	Indels	11	Gaps	5
QY	1	MSNRLLDGVAIIITGGTIGLAIATATKFEVEGAKVMYIDRHSDVEEKAASVOTPOIOFE	60						
DB	1	MSGRIGIGGVATVSGSGRGSGASHVRAMVAGEAKVFPFGDILDECKAVAAELA--DAARV	58						
QY	61	QHDSEDEGQWTLFDPATEKAFGPVSTLVNAGTAVNKSVEETTTAEKKLAVNLDGAF	120						
DB	59	HLDTQPRQMTAAVDTAVTAATVAFGLHVLVNNAGLNTICTIEDYALTEWRIIDLVNTLGV	118						
QY	121	CTRIGCIQMKKNGCLGASITNMSISIEGFVGDPSLGATVNASGCAVRIMSAAIDCAL	180						
DB	119	GIRAVAKPMKEKGRG-SITNISIEGLAGTVAACHGYATATKFAVAGLTFKSTALE--L	175						
QY	181	VRVTVTHGQYIKTPELVDDPGAAEAMASQRTKTCMGHITGCEPNDIVICYICVILAS	240						
DB	176	IRVNSIHGGLVTKPTMTDWP--EDIF-----QTALGRAAEVEVSNLVVYLLASDES	229						
QY	241	SEFVVDGCTYA	251						
DB	230	AEFVVDGCTVA	240						

RESULT 3  
S10707  
20beta-hydroxysteroid dehydrogenase (EC 1.1.1.-) - Streptomyces exfoliatus  
C.Species: Streptomyces exfoliatus  
C.Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C.Accession: S10707  
R.Marekov, L.; Kirook, M.; Joernvall, H.  
FEBS Lett. 266, 51-54, 1990  
A.Title: Prokaryotic 20-beta-hydroxysteroid dehydrogenase is an enzyme of the 'short-chain  
A.Reference number: S10707, MUID:90506362; PMID:2194840  
A.Accession: S10707  
A.Status: preliminary  
A.Molecule type: protein  
A.Residues: 1-255 <MAR>  
C.Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
C.Keywords: NAD: oxidoreductase  
F.7-183/Domains: short-chain alcohol dehydrogenase homology <SADH>

Query Match	31.6%	Score 407.5;	DB 1;	Length 255;
Best Local Similarity	39.0%;	Pred. No. 3,9e-25;		
Matches	97;	Conservative 42;	Mismatches 103;	Indels 7; Gaps 5

  

DB	3	NRLDGGVAITGTGLGIAIATKFEVGEAGVMITDRSDSGEKAASVTGPDOIQFOH	62
DY	1	:   :         :   :         :	
DY	2	NDLSGTIVITTGARGIGAEAKARQAVAAAGAVLADVLDEGSAATAARELG--DAARYOHL	59
DB	2	NDLSGTVITTTGARGIGAEAKARQAVAAAGAVLADVLDEGSAATAARELG--DAARYOHL	59
DY	63	DSSDEDGWTKLPDATEKAFGPVSTLVNNAAGIAVKKSVSEETTAERKLAVLVLDGFEST	122
DY	1	:   :   :   :   :             :   :	
DY	60	DVTIEEDMGGVAVAYAREPFGSVOSLVNNAAGISTGMLETFESVERRRKRVVDINLTVCFIGM	119

QY	123	RLGIDRMNRKNGKAGAIINMSSTIEGVOPSGAVNANAGVIRMSKAAIDCALDKYDVR	1822
	120	KTVIPAMMDAG-GGSIIVMISSAAGLGMALGLTSSIGASKMGYKGLSKLANVE--LGTDRIR	1767
QY	183	VNTHVHGIRKPLVLDPLPGAEELMSQRKTPMGHG-EPNDIATICYLASSNESKEATGS	2411
DB	177	VNSVHFGHTTYTPMAAE--TGIRQEGCNTPNTPMGVNGNEPEGLAGAVYKLLSDTSSYYTGA	2353
QY	242	EEVYVDGGTT	250
DB	236	ELAVDGGWT	244

RESULT 4  
S39737  
glucose 1-dehydrogenase homolog ywfD - *Bacillus subtilis*  
N:Alternate names: protein lpa-82d  
N:Contains: probable dehydrogenase (EC 1.1.1.-)  
C:Species: *Bacillus subtilis*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: S39737; E70055  
R:Glaser, P.; Kunst, F.; Arnard, M.; Coudart, M.F.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Repoport, G.; Danchin, A.  
Mol. Microbiol. 10, 371-384, 1993  
A:Title: *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region  
A:Reference number: S39655; MUID:95020537; PMID:794828  
A:Accession: S39737  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-555 <GLN>  
A:Cross-references: EMBL:773124; NID:g413923; PIDN:CA51638.1; PID:g414006  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.: Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Caplan, V.; Carter, N.M.;  
A.: Ehrlich, S.D.; Emmerson, P.T.; Eutani, K.D.; Errington, J.; Fabret, C.; Ferrarri,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, C.R.; Hentz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
lech, J.; Harwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidos, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Me  
Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tepstorf, P.; Tognoni, A.; Tosato, Y.; Uchida  
T.; Winters, P.; Wipak, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danclon, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtili*  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: E70055  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-255 <KUN>  
A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15799.1; PID:g26363  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yw4D  
C:Superfamily: ribitol dehydrogenase: short-chain alcohol dehydrogenase homology  
C:Keywords: NAD; oxidoreductase  
F:8-185/Domain: short-chain alcohol dehydrogenase homology <SADH>

	Query Match	31.0%	Score 400.5;	DB 1:	Length 255;
	Best Local Similarity	37.1%;	Pred. No. 1.4e-24;		
Matches	96;	Conservative	45;	Mismatches	107; Indels 11; Gaps
OY	1 MSNRLDGKVAIIITGGTGLIGLAIAIRFVEBGAKYMITDRHSDVGEKAASKSGPDIQIF 60				
			:	:	:
Db	1 MIMNLDTKVLTITGASGICGYAAVFLGQAQAVVVADIDEOGEAAYRKENN-DRLHFV 59				
			:	:	:
OY	61 QHDSDEDCGTTKLEPDATETAKAEFPVSFLVNAGIAYVNSVEETTAAEMRKLLAVNLDOVF 120				
			:	:	:
Db	60 QYDITDEACQAHAVEASVHFEGGLDVILINNAIEIYAPIHMEIELSDNNKYQLQYNLTGMFL 119				
			:	:	:
OY	121: GTPLGIQRKKNKKLGAJIIMSSIEGFVGDPISGAIYNAKGAVRIINSKSAALDCAIKDYD 180				



DB 120 MSKHALKMLAAKRG-NIIMTCSVGLVAMPDIPAYNASKGVQLTKRMAVDYA--KHQ 176  
 QY 181 VRVNVHPGYIKTPPL-----VDDLPGA-EEAMSORIK-TPMGHIGEPNDIAYICVYLAASN 233  
 DB 177 IRVNCVCPGIIDPILNKSKFLENNEGTLLEIKREKAKVNPILRRGKREELANVWLPLASD 236  
 QY 234 ESKFATGSEFVVOGGYTAQ 252  
 DB 237 LSSYMTGSAITADGGYTAQ 255

## RESULT 5

3(or 17)beta-hydroxysteroid dehydrogenase (EC 1.1.1.51) - Comamonas testosteroni (ATCC 154129)  
 C:Species: Comamonas testosteroni  
 A:Variety: ATCC 11996  
 C:Date: 14-Jul-1995 #sequence, revision 01-Dec-1995 #text\_change 08-Oct-1999  
 C:Accession: S48129; S51780; S15390; S62216; S62182  
 R:Abalain, J.H.; di Stefano, S.; Amet, Y.; Quemener, E.; Abalain-Colloc, M.L.; Floch, H.  
 J. Steroid Biochem. Mol. Biol. 44, 133-139, 1993  
 A:Title: Cloning, DNA sequencing and expression of (3-17)beta hydroxysteroid dehydrogenase  
 A:Reference number: S48129; MUID:93176721; PMID:8382516  
 A:Accession: S48129  
 A:Molecule type: DNA  
 A:Residues: 1-254 <BA>  
 A:Cross-references: EMBL:X63379  
 A:Note: The source is designated as Pseudomonas testosteroni  
 R:Abalain, J.H.  
 submitted to the EMBL Data Library, November 1991  
 A:Reference number: S51780  
 A:Accession: S51780  
 A:Molecule type: DNA  
 A:Residues: 1-13, 'VV', 16-254 <AB>  
 A:Cross-references: EMBL:X63379; NID:9312918; PIDN:CAA44977.1; PID:9312919  
 A:Note: The source is designated as Pseudomonas testosteroni  
 R:Yin, S.J.; Vagelopoulos, N.; Lundquist, G.; Joernvall, H.  
 Eur. J. Biochem. 197, 359-365, 1991  
 A:Title: Pseudomonas 3-beta-hydroxysteroid dehydrogenase. Primary structure and relation  
 A:Reference number: S15390; MUID:91224137; PMID:2026158  
 A:Accession: S15390  
 A:Molecule type: protein  
 A:Residues: 2-40, 'E', 41-176, 178-240, 'G', 242-254 <YN>  
 A:Note: The source is designated as Pseudomonas testosteroni  
 R:Bench, J.; Knapp, S.; Oppermann, U.C.T.; Haegglund, O.; Joernvall, H.; Ladenstein, R.  
 Eur. J. Biochem. 236, 144-148, 1996  
 A:Title: Crystallization and crystal packing of recombinant 3 (or 17) beta-hydroxysteroid  
 A:Reference number: S62216; MUID:96184891; PMID:8617258  
 A:Accession: S62216  
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 2-254 <BN>  
 R:Bocco, J.L.; Panzetta, G.; Actis, L.A.; Gentl-Raimondi, S.  
 submitted to the EMBL Data Library, July 1993  
 A:Description: Nucleotide sequence of the beta-hydroxysteroid dehydrogenase from Pseudom  
 A:Reference number: S62182  
 A:Accession: S62182  
 A:Molecule type: DNA  
 A:Residues: 1-40, 'E', 41-176, 178-254 <BO>  
 A:Cross-references: EMBL:L08971; NID:9309859; PIDN:AA25742.1; PID:9309860  
 A:Note: The source is designated as Pseudomonas testosteroni  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C:Keywords: oxidoreductase  
 F:2-354/Product: 3(or 17)beta-hydroxysteroid dehydrogenase #status experimental <MAT>  
 F:8-185/Domain: short-chain alcohol dehydrogenase homology <SABH>

## Query Match

Best Local Similarity 30.5%; Score 393; DB 2; Length 254;  
 Matches 92; Conservative 49; Mismatches 100; Indels 10; Gaps 6;

QY 1 MSNRDQKVAITGTGIGLAITATFVEEGAVMITDRHSVDGEKAKSVGTPDOIOFP 60  
 DB 1 MTRNLQKVALVYGASVGLLEVKKLLIGGAKAVAFSDINA-AGQOLAELG--ERSMF 57

QY 61 QHSDSDGCKTLKEDATEKAGPSTLVNNGIAVNKSVEETTTAEWRKLLAVNLDCVF 120  
 DB 58 RHVDSSEADWTLVMAAORLGLTLNVLNNAIGILLPGDMETGRLEDSRLKITEVFI 117  
 QY 121 CTRGLQRMKRGKIGASTINNSSTEGYDPSLCAYNASKAVINMSKALDCAKLDYD 180  
 DB 118 GCOOGIAAMKE--TGGSTINNASVSMPLPEOYAGYSKSKAAVSALTAAALSCRKGYA 175  
 QY 181 V-RVNVHPGYIKTPPLVD-DLP--GAEEAMSORIK-TPMGHIGEPNDIAYICVYLAASN 235  
 DB 176 IRRVNSIHDPGIIITPMQASLPKGVSKKENVLHDKRLNRAGRAVPERIALVLFASDES 235  
 QY 236 KEATGSEFVVD 246  
 DB 236 SVMSGSELHAD 246

## RESULT 6

547055  
 hypothetical protein 5 - Xanthobacter sp.  
 C:Species: Xanthobacter sp.  
 C:Date: 19-Mar-1997 #sequence, revision 19-Mar-1997 #text\_change 05-Nov-1999  
 C:Accession: S47055  
 R:Swaving, J.; Weijers, C.A.G.M.; van Ooyen, A.J.J.; de Bont, J.A.M.  
 submitted to the EMBL Data Library, June 1994  
 A:Description: Plasmidation of Xanthobacter Py2 mutants in epoxylane degradation;  
 A:Reference number: S47051  
 A:Accession: S47055  
 A:Molecule type: DNA  
 A:Residues: 1-249 <SW>  
 A:Cross-references: EMBL:X79863; NID:9520947; PIDN:CAA56245.1; PID:9520952  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 F:5-181/Domain: short-chain alcohol dehydrogenase homology <SABH>

## Query Match

Best Local Similarity 30.3%; Score 391; DB 2; Length 249;  
 Matches 98; Conservative 39; Mismatches 105; Indels 10; Gaps 6;

QY 5 LDGKVAITGTGIGLAITATFVEEGAKVMITDRHSDVEKAKSVGTPOIOFQHDS 64  
 DB 2 LDEVAITGTGAAGIGLAVHAARAGARVALLDRDCAQORAAEFGA--AMGVADY 59  
 QY 65 SDEGCKTLKEDATEKAGPSTLVNNGIAVNKSVEETTTAEWRKLLAVNLDCVF 124  
 DB 60 TDEAIIAAMAGQORALGPIGLVNNAGIAGFGSVHATEVETWSRIMAVVTGTFASKA 119  
 QY 125 GIDRMKKGIGASTINNSSTEGYDPSLCAYNASKAVINMSKALDCAKLDYDVRN 184  
 DB 120 ALGEMLERGKA-IVNGSVAGLVGIPYMAATCAKAAVNLTRQMAVDSGR--GIRVN 176  
 QY 185 TVHPGYIK-TPLVDDLPGE--EAMSOR-TKTPMGHIGEPNDIAYICVYLAASNESKFAT 239  
 DB 177 VVCGTGTAGIDMGKQLGCTDCELEKRLAKTPMGFGTIPEDIAEAVALFLSTKAFTV 236  
 QY 240 GSEFVVDGYTA 251  
 DB 237 GSVLAVDGGMTA 248

## RESULT 7

E72427  
 oxidoreductase, short chain dehydrogenase/reductase family - Thermotoga maritima (s  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence, revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: E72427  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; H  
 Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genom  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: E72427

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-256 <ARN>  
 A:Cross-references: GB:AE001690; GB:AE000512; NID:94980456; PIDN:ADJ35113.1; PID:9498050  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: JH0019  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 29.7%; Score 383; DB 2; Length 256;  
 Best Local Similarity 35.5%; Pred. No. 3.5e-23;  
 Matches 93; Conservative 48; Mismatches 95; Indels 26; Gaps 7;

QY 5 LQGVAVITGCTGIGLAIATKFEVEGAKVMTDRHSDVC---EKAASVGPDPQIOFQ 61  
 DB 2 LQGVAVITGCTGIGLAIATKFEVEGAKVMTDRHSDVC---EKAASVGPDPQIOFQ 59  
 QY 62 HDSDDGKTKLEDATEKAPSTLVNNAAGIANKSVSEETTTAEWRKLLAVNLDGVFQ 121  
 DB 60 TDVADNSVKNMVKRYETVGVVAVNNAVSVKSIPEERLEMERVIRVNLGPFYIC 119  
 QY 122 TRAGIORMKNKKGASIIINMSIEGFVGPDSLGAVNASKAVNMSKALDCAKDYD 181  
 DB 120 SNVCAEEMKRG-GVYIINASTRAFOSEPDPTEPSKSGGLVALTHSLAV--SLSRVHI 176  
 QY 182 RNVNHPGTYIKTPVLDLPGAEFAMQRTK-----PMGHIGEPNDIAYICVYLA 231  
 DB 177 RVVVISPGIET-----SEMKRKSIRKKRDLRPIDHQHPAGRGVGNPLDIHLCEFLA 229  
 QY 232 SNE-SKFAFGSEFVVDGCTAQ 252  
 DB 230 DDEKAGFTGTNTFVDDGGMVTK 251

# RESULT 8

Probable fabG2 protein - Mycobacterium tuberculosis (strain H37RV)  
 E70740

C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: E70740

R:Coile, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Brown, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: E70740

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-247 <COL>

A:Cross-references: GB:275555; GB:AL123456; NID:93261608; PIDN:CAA9983.1; PID:91419053  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: fabG2

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 F:8-186/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 29.5%; Score 380; DB 2; Length 247;  
 Best Local Similarity 36.2%; Pred. No. 5.8e-23;  
 Matches 92; Conservative 52; Mismatches 94; Indels 16; Gaps 4;

QY 1 MSNRLOGVAITGCTGIGLAIATKFEVEGAKVMTDRHSDVGEKAKSVGTPQIOFQ 60  
 DB 1 MASLIMARTAVITGCGAGLAIQGRFAEGARVLDGVNLEATEVAARKRLGDDVAAV 60  
 QY 61 QHDSDEGKTKLEDATEKAPSTLVNNAAGIANKSVSEETTTAEWRKLLAVNLDGVF 120  
 DB 61 RCDVQADVDVILRTAVAREGGLDVVNVNAGITRDATRTTEBQFOQVIAVHLKGTWN 120  
 QY 121 GTRAGIORMKNKKGASIIINMSIEGFVGPDSLGAVNASKAVNMSKALDCAKDYD 180  
 DB 121 GTRAGIORMKNKKGASIIINMSIEGFVGPDSLGAVNASKAVNMSKALDCAKDYD 180

QY 181 VRVNVHPGTYIKTPVLDLPGAEFAMQRTK-----TKTPMGHIGEPNDIAYICVYLA 234  
 DB 178 IRVNAIAGGLIRSAMT-----EAMPQRIHQDLAEVPMGRAGESEVAVAVFLASDL 230  
 QY 235 SKFATGSEFVVDG 248  
 DB 231 SYMTGTVLDVYTG 244

# RESULT 9

Glucose 1-dehydrogenase homolog ykwo - Bacillus subtilis  
 F69868

C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: F69868

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Enliam, K.D.; Ertlington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lartino Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekova, A.; Se akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchida T.; Winters, P.; Wipal, A.; Yamamoto, H.; Yamano, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F69868

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-248 <RUN>

A:Cross-references: GB:299111; GB:AL009126; NID:92633699; PIDN:CAJ13250.1; PID:926337  
 A:Experimental source: strain 168  
 C:Genetics:

A:Gene: ykwo  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 F:7-182/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 29.0%; Score 374.5; DB 2; Length 248;  
 Best Local Similarity 37.5%; Pred. No. 1.6e-22;  
 Matches 93; Conservative 40; Mismatches 106; Indels 9; Gaps 5;

QY 4 RLQGVAVITGCTGIGLAIATKFEVEGAKVMTDRHSDVGEKAKSVGTPQIOFQ 63  
 DB 3 KFEKIALVYTCGTSIGLAIATKFEVEGAKVMTDRHSDVGEKAKSVGTPQIOFQ 60  
 QY 64 SSODEGKTKLEDATEKAPSTLVNNAAGIANKSVSEETTTAEWRKLLAVNLDGVFQ 123  
 DB 61 ISKLEDEKTKLEDATEKAPSTLVNNAAGIANKSVSEETTTAEWRKLLAVNLDGVFQ 120  
 QY 124 LGIORMKNKKGASIIINMSIEGFVGPDSLGAVNASKAVNMSKALDCAKDYD 183  
 DB 121 KALSTPEPK--VGSITVGTAGSICGNPAPSVGASRAALRAVLRNML--LGTEIRV 176  
 QY 184 NVNHPGTYIKTPVLDLPG--AEBAM--SQRTKTPMGHIGEPNDIAYICVYLA 240  
 DB 177 NVNHPGTYIKTPVLDLPG--AEBAM--SQRTKTPMGHIGEPNDIAYICVYLA 240  
 QY 241 SEFVVDG 248  
 DB 237 VELFVDG 244

# RESULT 10

Probable (Imported) - Sinorhizobium meliloti (strain 1021) magaplasmid psyma  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: B95284  
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
 K.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
 Proc. Natl. Acad. Sci. U.S.A. 99, 9883-9888, 2001  
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*  
 A:Reference number: A95262; MUID:21396509; PMID:11481432  
 A:Accession: B95284  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-253 <KUR>  
 A:Cross-references: GB:AE006469; PIDN:AAK64836.1; PID:g14523249; GSPDB:GN00165  
 A:Experimental source: strain 1021, megaplasmid pSymA  
 R:Galibert, F.; Flman, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Sma0329  
 C:Genome: Plasmid  
 A:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 28.9%; Score 372.5; DB 2; Length 253;  
 Best Local Similarity 37.7%; Pred. No. 2.4e-22;  
 Matches 97; Conservative 46; Mismatches 93; Indels 21; Gaps 8;

OY 1 MSNRDGVVAITGCTGIGLAITATKFEVSGAKVMTDRH-----SDVGEKRAKSVGTPDQ 56  
 1 MSKRFDCVVAITVGGSGGCAIATNLEEGASVMSGRTKRLSDVASKM-----PADR 55  
 DB 1 MSKRFDCVVAITVGGSGGCAIATNLEEGASVMSGRTKRLSDVASKM-----PADR 55

OY 57 IQOFHDSDEDEGWTKLDATEKAFQPVSTLVNAGIAVNSVEETTAEMRKLAVNLD 116  
 DB 56 SGIRFVANSRDCALVAATVERFGRIDTVNAGNFEVGTIOETSDQDMDIASDLS 115

OY 117 GVFEGTRIGIORMK-NKGIGASTINMSIEGFQDPSLGAVNASKGAVRINSKAAALDCA 175  
 DB 116 GVFVMSRAVAPLTKETK--SIVNIGSVSSLGGMASHAAYNAKGVAMLTRSAA--CD 170

OY 176 LKDYDVAVNTVHPGYIKTPLVDLPAGEAMSGRT--KTPNGHIGEPNDIAYICVYLASN 233  
 DB 171 LKGFGRANVTAPGLTGVGMVAIMD-DDALLEKANDRIPLRRAGQP--ASAVAFILASD 226

OY 234 ESKFATGSEFVVDGGYT 250  
 DB 227 EAAMITGIVLPVDCGT 243

RESULT 11  
 JS0385  
 glucose 1-dehydrogenase (EC 1.1.1.47) - *Bacillus megaterium*  
 C:Species: *Bacillus megaterium*  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 11-Jun-1999  
 A:Accession: J39852; JS0385  
 R:Matsuura, T.; Eboru, R.Y.; Nakai, T.; Makino, Y.; Negoro, S.; Urabe, I.; Okada, H.  
 J. Ferment. Bioeng. 70, 363-369, 1990  
 A:Title: Structure of isozyme genes of glucose dehydrogenase from *Bacillus megaterium* IA  
 A:Reference number: J39850  
 A:Accession: J39852  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-261 <RES>  
 A:Cross-references: GB:D90043; NID:g216265; PIDN:BAAL4099.1; PID:g216268  
 A:Experimental source: strain IAM1030  
 C:Complex: homotetramer  
 C:Function:  
 A:Description: catalyzes NAD(P)-dependent dehydrogenation of D-glucose.  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C:Keywords: oxidoreductase  
 F:8-189/Domain: short-chain alcohol dehydrogenase homology <SAND>

Query Match 28.8%; Score 371.5; DB 2; Length 261;  
 Best Local Similarity 34.2%; Pred. No. 3e-22;  
 Matches 88; Conservative 51; Mismatches 105; Indels 13; Gaps 5;

OY 1 MSNRDGVVAITGCTGIGLAITATKFEVSGAKVMTDRHSDVGEKA-----ASVGP 54  
 1 MKDLEGKVVVITGSSGGLKAMAIRATEKAKVAVNRSKE--EANSVLEIKVKG-- 56  
 DB 1 MKDLEGKVVVITGSSGGLKAMAIRATEKAKVAVNRSKE--EANSVLEIKVKG-- 56

OY 55 DQIOFHDSDDEGWTKLDATEKAFQPVSTLVNAGIAVNSVEETTAEMRKLAVN 114  
 DB 57 GEIAVAKGDVTESDVINTVOSAIKEFKGLDVMINMAGMENPVSHEMSLSDMNKVIIDN 116

OY 115 LQGFEGTRIGIORMKKNKGIGASTINMSIEGFQDPSLGAVNASKGAVRINSKAAALDC 174  
 DB 117 LTFGFLGSRRAIKFYVENDIKGTVINMSVHEKIPWPLFVHYAASKGKMLTETLALRY 176

OY 175 ALKDYDVAVNTVHPGYIKTPL-VDDLPAGEAMSGRTTPNGHIGEPNDIAYICVYLASN 233  
 DB 177 AFR--GIRVNNIGPAINTPINAEKFADEQRAVDESMPGVIIGEPPEIAVAAMVLASS 234

OY 234 ESKFATGSEFVVDGGYT 250  
 DB 235 EASYVTGITLFDAGCMT 251

RESULT 12  
 A33528  
 glucose 1-dehydrogenase (EC 1.1.1.47) - *Bacillus megaterium*  
 C:Species: *Bacillus megaterium*  
 C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 11-Jun-1999  
 A:Accession: A33528  
 R:Makino, Y.; Negoro, S.; Urabe, I.; Okada, H.  
 J. Biol. Chem. 264, 6381-6385, 1989  
 A:Title: Stability-increasing mutants of glucose dehydrogenase from *Bacillus megate*  
 A:Reference number: A33528; MUID:89197943; PMID:2495285  
 A:Accession: A33528  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-261 <MAK>  
 A:Cross-references: GB:J04805; NID:g142974; PIDN:AAA22475.1; PID:g142975  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C:Keywords: oxidoreductase  
 F:8-189/Domain: short-chain alcohol dehydrogenase homology <SAND>

Query Match 28.7%; Score 370.5; DB 2; Length 261;  
 Best Local Similarity 33.6%; Pred. No. 3.6e-22;  
 Matches 86; Conservative 51; Mismatches 108; Indels 11; Gaps 4;

OY 1 MSNRDGVVAITGCTGIGLAITATKFEVSGAKVMT-----DRHSDVGEKRAKSVGTPD 55  
 DB 1 MYKDLGKVVVITGSSGGLKSMAIRATEKAKVAVNRSKEDANSVLEIKVKGCAI 60

OY 56 QIOFHDSDDEGWTKLDATEKAFQPVSTLVNAGIAVNSVEETTAEMRKLAVN 115  
 DB 61 AV---KGDVTESDVINTVOSAIKEFKGLDVMINMAGMENPVSHEMSLSDMNKVIIDN 117

OY 116 DGVFEGTRIGIORMKKNKGIGASTINMSIEGFQDPSLGAVNASKGAVRINSKAAALDCA 175  
 DB 118 TGAFGLSREAIKFEVENDIKGTVINMSVHEKIPWPLFVHYAASKGKMLTETLALRY 177

OY 176 LKDYDVAVNTVHPGYIKTPL-VDDLPAGEAMSGRTTPNGHIGEPNDIAYICVYLASN 234  
 DB 178 PK--GIRVNNIGPAINTPINAEKFADEQRAVDESMPGVIIGEPPEIAVAAMVLASS 235

OY 235 SKFATGSEFVVDGGYT 250  
 DB 236 ASYVTGITLFDAGCMT 251

RESULT 13  
 B86737  
 acetoin dehydrogenase (EC 1.1.1.5) [imported] - *Lactococcus lactis* subsp. *lactis* (s

RESULT 14  
C70885  
probable dehydrogenase - *Mycobacterium tuberculosis* (strain H37RV)  
C:Species: *Mycobacterium tuberculosis*  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: C70885  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtzoy, R.; Janda, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: C70885  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-258 <COL>  
A:Cross-references: GB:AL008883; GB:AL123456; NID:93261490; PIDN:CA115519.1; PID:9261286  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV2857c  
C:Superfamily: rbtitol dehydrogenase; short-chain alcohol dehydrogenase homology <SAOH>  
F:11-188/Domain: short-chain alcohol dehydrogenase homology <SAOH>

Query Match	28.48;	Score 367;	DB 2;	Length 258;
Best Local Similarly	35.7%;	Pred. No. 6,7e-22;		
Matches 92;	Conservative 45;	Mismatches 105;	Indels 16;	Gaps 7;
QY	1	MSNRDQKVALITGCTLGIGLAIATKFEYEEAKWITDRHSDVGEKAAKSVGTPOIQ--	58	

**RESULT 15**  
 A99950  
 hypothetical protein SA2260 [Imported] - *Staphylococcus aureus* (strain N315)  
 C:Species: *Staphylococcus aureus*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: A99950  
 R:Kurodo, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, ma, A.; Mutant-01, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimura, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hitamatsu, K.  
 L:ncnt 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: A99950  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-272 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13702422; PIDN:BA843563.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SA2260  
 A:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homolog

[illegible]

Search completed: April 22, 2003, 17:16:52  
Job time : 22 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 17:12:56 ; Search time 13 Seconds  
(without alignments)  
804.003 Million cell updates/sec

Title: US-09-910-033A-2

Perfect score: 1290  
Sequence: 1 MSNRDQKVAITTCGLIG.....NESKPAFGSEFVVDGGYTAQ 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	464.5	36.0	260	1 YK02_MYCTU	Q10855 pseudocactari
2	441.5	34.2	250	1 LINX_PSEPA	P50198 pseudomonas
3	407.5	31.6	255	1 2BHD_STREX	P19992 streptomyc
4	400.5	31.0	255	1 YNFD_BACSU	P39640 bacillus su
5	383	29.7	256	1 Y019_THEMA	Q56318 thermotoga
6	380	29.5	247	1 YD50_MYCTU	Q11020 mycobacteri
7	371.5	28.8	261	1 DHG1_BACME	P39482 bacillus me
8	370.5	28.7	253	1 DHG_BACME	P40288 bacillus me
9	370	28.7	253	1 3BHD_COMTE	P19871 comamonas t
10	365.5	28.3	246	1 FABG_THEMA	Q9X248 thermotoga
11	363.5	28.2	261	1 DHGA_BACME	P10528 bacillus me
12	360	27.9	262	1 DHGB_BACME	P07999 bacillus me
13	358.5	27.8	261	1 DHG_BACSU	P12310 bacillus su
14	357.5	27.7	261	1 DHG4_BACME	P39485 bacillus me
15	355.5	27.6	261	1 DHG2_BACME	P39483 bacillus me
16	352.5	27.5	261	1 DHG3_BACME	P39484 bacillus me
17	351.5	27.2	251	1 Y325_THEMA	Q9WY90 thermotoga
18	348.5	27.0	258	1 DHG2_BACSU	P08089 bacillus su
19	345	26.7	261	1 DHB8_HUMAN	Q92506 homo sapien
20	339.5	26.3	256	1 BUDC_KLEPN	P08436 klebsiella
21	326	25.3	285	1 GS39_BACSU	P80893 bacillus su
22	325.5	25.2	260	1 DHB8_MOUSE	P50171 mus musculu
23	324	25.1	336	1 T52_MAIZE	P50160 zea mays (m
24	322	25.0	289	1 YHDF_BACSU	Q00750 bacillus su
25	321	24.9	248	1 FABG_CHLNU	Q09K77 chlamydia s
26	321	24.9	248	1 Y4VI_RHISN	Q53217 rhizobium s
27	319.5	24.8	256	1 GNO_GLUOX	P50199 gluconobact
28	319.5	24.8	258	1 BRHA_RHIME	Q86034 rhizobium m
29	318	24.7	244	1 FABG_VIBCH	Q9K977 vibrio chol
30	317.5	24.6	262	1 VER1_ASPPA	P50161 aspergillus
31	317	24.6	241	1 PHBB_ZOORA	P23238 zoogloea ra
32	316	24.5	246	1 FABG_BACSU	P51831 bacillus su
33	316	24.5	281	1 SOUL_CANAL	P87219 candida alb

34	313.5	24.3	264	1 STCU_EMENT	Q00791 emeticella
35	312	24.2	241	1 PHBB_RHIME	P50205 rhizobium m
36	312	24.2	248	1 FABG_AQUAE	Q67610 aquifex aeo
37	310.5	24.1	271	1 TRN1_PICAB	Q08632 picea abies
38	310	24.0	273	1 TRN1_DARTS	P50197 datura stra
39	307.5	23.8	250	1 LINX_PSEPA	P50197 pseudomonas
40	305	23.6	247	1 FABG_CHLFR	P38004 chlamydia t
41	305	23.6	285	1 YHXC_BACSU	P40397 bacillus su
42	304.5	23.6	248	1 FABG_CHLNU	Q948P2 chlamydia s
43	303	23.5	278	1 Y4LA_RHISN	P55541 rhizobium s
44	301.5	23.4	320	1 FABG_CUPLA	P28643 cuphea lanc
45	299.5	23.2	253	1 Y4MP_RHISN	P55575 rhizobium s

## ALIGNMENTS

RESULT 1	ID	YK02_MYCTU	STANDARD	PRT	260 AA.
AC	Q10855:				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	15-JUL-1999 (Rel. 38, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Putative oxidoreductase Rv2002 (EC 1.-.-.-)				
GN	FABG3 OR Rv2002 OR MT2058 OR MTCY39.16C.				
OS	Mycobacterium tuberculosis.				
OC	Bacteria: Actinobacteria; Actinobacteria (class); Actinobacteridae;				
OC	Actinomycetales: Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1773;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-H37RV;				
RX	MEDLINE=98295987; PubMed=9634230;				
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,				
RA	Gordon S.V., Eigler K., Gas S., Barry C.E. III, Telata F.,				
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,				
RA	Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,				
RA	Hornsby T., Jagels K., Krogan A., McLean J., Moule S., Murphy L.,				
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,				
RA	Rutler S., Seeger K., Skelton S., Squares S., Squares R.,				
RA	Sulton J.E., Taylor K., Whitehead S., Barrell B.G.,				
RA	"Deciphering the biology of Mycobacterium tuberculosis from the				
RT	complete genome sequence."				
RT	Nature 393:537-544(1998).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-CDC 1551 / Oshkosh;				
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,				
RA	Petersen J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,				
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,				
RA	Delcher A., Utterback T., Weidman J., Khoult H., Gill J., Mikule A.,				
RA	Bisai W.;				
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and				
RT	laboratory strains."				
RL	Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.				
CC	-I- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES				
CC	(SDR) FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	EMBL: Z74025; CAA98414.1; -				
DR	EMBL: AE007057; AAK46335.1; -				
DR	HSNP: P19992; IHDC.				
DR	TIGR: MT2058; -				
DR	Tuberculist: Rv2002; -				
DR	InterPro: IPR002198; ADH_short.				

DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT; 1.  
 KW Hypothetical protein: Oxidoreductase: NAD: Complete proteome.  
 FT ACT\_SITE 153 153 NAD (BY SIMILARITY).  
 FT BIND 174 174 BY SIMILARITY.  
 FT CONFLICT 174 174 S -> G (IN REF. 2).  
 SQ SEQUENCE 260 AA; 27030 MW; 0935A14ED36220B7 CRC64;

Query Match 36.0%; Score 464.5; DB 1; Length 260;  
 Best Local Similarity 42.2%; Pred. No. 6,6e-31;  
 Matches 106; Conservative 44; Mismatches 90; Indels 11; Gaps 5;

OY 1 MSNRDGLKVAITTGTLGIGLAIAATKEVEGAKVMITDRHSVDGEKAASVGPDPDIOFE 60  
 DB 1 MSGRILGKVALVSGARGSGASHVAMVEGAKVFGDLDEBKAVALAELA--DAARIV 58  
 OY 61 QHDSDEDCGWTLPDTEAKFPGVSTLVNAGIAVKNKSVETTTAEWRKLLAVNLGVFF 120  
 DB 59 HLDVTPQAWMTAVDPYAVTAFFGLHLVNNAGILNIGTIEDYALTEMORILLDNLGVFL 118  
 OY 121 GTRIGIQRKKNKGLASIIINMSIEGFGVDPISLGAIVNASKAVRIMSKAALDCAIKDYD 180  
 DB 119 GTRAVVKKPKKKEGRG--SIINISIEGLACTVACHGTATKFAVRKGLTKSTALE--LGPSG 175  
 OY 161 VRVNTVHPGVIKTPVLDLPGAEAMSORTKTPMGHIGEPNDIAYICVYLASNESEKATG 240  
 DB 176 IRVNSIHGIVTPTMDVP--EDIF---GTALGRAAEVPEVSNLYVLADSESYISG 229  
 OY 241 SEFVVDGCVTA 251  
 DB 230 AEFVVDGCVTA 240

## RESULT 2

LINK\_PSEPA  
 ID LINK\_PSEPA STANDARD: PRT: 250 AA.

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase (EC 1.1.-.-)  
 GN (2,5-DOOL dehydrogenase).

OS Pseudomonas paucimobillis (Sphingomonas paucimobillis).  
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
 CC Sphingomonas.

NCBI\_TaxID=13689;

OX 11

RP SEQUENCE FROM N.A.

RC STRAIN=UT26;

RX MEDLINE=94252977; PubMed=7515041;

RA Nagata Y., Ohtomo R., Miyachi K., Fukuda M., Yano K., Takagi M.;

RT Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol

RT dehydrogenase gene involved in the degradation of gamma-

RT hexachlorocyclohexane in Pseudomonas paucimobillis.\*;

RL J. Bacteriol. 176:3117-3125(1994).

-1- FUNCTION: DEGRADATION OF 2,5-DICHLORO-2,5-CYCLOHEXADIENE-1,4-DIOL

(2,5-DDOL) INTO 2,5-DICHLOROHYDROQUINONE (2,5-DCHQ). LINK IS NOT

ESSENTIAL TO GAMMA-HCH DEGRADATION.

-1- PATHWAY: DEGRADATION OF GAMMA-HEXACHLOROCYCLOHEXANE.

-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

(SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D23722; BAA04939.1; -

DR HSPR: P19992; IHNC.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase: NAD.  
 FT ACT\_SITE 156 156 NAD (BY SIMILARITY).  
 FT BIND 174 174 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 25492 MW; 8C52703FF76382CF CRC64;

Query Match 34.2%; Score 441.5; DB 1; Length 250;  
 Best Local Similarity 41.0%; Pred. No. 4,7e-29;  
 Matches 105; Conservative 43; Mismatches 93; Indels 15; Gaps 6;

OY 1 MSNRDGLKVAITTGTLGIGLAIAATKEVEGAKVMITDRHSVDGEKAASV-----GTPD 55  
 DB 1 MANRLAGKVALITGSGASGIAAOKRAFEAGAKVY----GDLNEEMAGVVAEIRIAGG 56  
 OY 56 QIQFOHSSDEDCGWTLPDTEAKFPGVSTLVNAGIAVKNKSVETTTAEWRKLLAVNL 115  
 DB 57 DALFTRLDVTDAASWNNALAAAVDGGGLTTLNNTAGIHPGCFEESIEGNNKVAANQ 116  
 OY 116 DGVFEGRIGIQRKKNKGLASIIINMSIEGFGVDPISLGAIVNASKAVRIMSKAALDCA 175  
 DB 117 TAFILGKAIPELVKSSNG--SIINISLIGMFPAGNASYCATRAAVIMSKAAL-- 173  
 OY 176 LKDYVRVNTVHPGVIKTPVLDLPGAEAMSORTKTPMGHIGEPNDIAYICVYLASNE 234  
 DB 174 FVDRGVRVNTVHPGGMNPTITANVP--PDVLKQRTSOLPMKLGIDPIDIANCALFLASE 231  
 OY 235 SKFATSEFVVDGCVT 250  
 DB 232 AKYITGVLDPLDGWS 247

## RESULT 3

2BHD\_STREX  
 ID 2BHD\_STREX STANDARD: PRT: 255 AA.

AC P19992;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE 20-beta-hydroxysteroid dehydrogenase (EC 1.1.1.53).

OS Streptomyces exfoliatus (Streptomyces hydropneumans).

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

CC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.

NCBI\_TaxID=1905;

OX 11

RP SEQUENCE.

RC MEDLINE=90306362; PubMed=2194840;

RA Ghosh D., Weeks C.M., Grochulski P., Duax W.L., Erman M.;

RT "Prokaryotic 20 beta-hydroxysteroid dehydrogenase is an enzyme of the

RT 'short-chain, non-metalloenzyme' alcohol dehydrogenase type.\*;

RL FEBS Lett. 266:51-54(1990).

RL 121

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).

RX MEDLINE=92052211; PubMed=1946424;

RA Ghosh D., Weeks C.M., Grochulski P., Duax W.L., Erman M.;

RT "Three-dimensional structure of holo 3 alpha,20 beta-hydroxysteroid

RT dehydrogenase: a member of a short-chain dehydrogenase family.\*;

RL Proc. Natl. Acad. Sci. U.S.A. 88:10064-10068(1991).

-1- CATALYTIC ACTIVITY: Androstane-3-alpha,17-beta-diol + NAD(+) -> 17-

-1- beta-hydroxyandrostane-3-one + NADH.

-1- SUBUNIT: HOMOTETRAMER.

-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

(SDR) FAMILY.

PIR: S10707; S10707.

PDB: 2HSD; 31-AUG-94.

PDB: 1HDC; 07-FEB-95.

InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short; 1.

DR PRINTS: PR00080; SDRFAMILY.

DR PROSITE: P500061; ADH\_SHORT; 1.  
 KW Oxidoreductase: NAD; Steroid metabolism; 3D-structure.  
 FT NP\_BIND 10 34 NAD (BY SIMILARITY).  
 FT ACT\_SITE 152 152  
 SO SEQUENCE 255 AA; 26484 MW; 9CB93CB66AA628D5 CRC64;

Query Match  
 Best Local Similarity 31.0%; Score 407.5; DB 1; Length 255;  
 Matches 97; Conservative 42; Mismatches 103; Indels 7; Gaps 5;

DB 3 NRLLDQKVAITGGTGLGIAATKFEVEGAKVMTDRHSVDGEKAASVGTPOIOFROH 62  
 2 NLSGCTVITITGARGLGAEARQAVAGARVLAADVDEGATARELG--DAARQHL 59  
 63 DSSDEDGWTKLDATEKAFGPVSTLVNAGIAVANKSVEETTAEMRKLLAVNLGCVFET 122  
 60 DVTEEDMORVAVAYAREEGSGVDGLVNNAGISITGMEFETESVERFRKVDINTLTCVFIGM 119  
 123 RLGIOMKMKKGGASTINNSIEGFPVDSGLAVNASKGAVRIMSKAALCALKDQYVR 182  
 120 KTVIPPMKAG--GGSTVNISSAAGLWGLALTSSYGASKMGVRLSKLAAYE--LCTDRIR 176  
 183 VNTVHPGYIKTFLVDDLPAGEAASQRTKTPMGHIG-EPNDIAVICYVLASNESKFAFGS 241  
 177 VASVHPGMYTMTAE--TGIRGEGCNVPTPMGRVGNBPETIAGAVYKLLSDTSYVTGA 235  
 242 EFVVDGQY 250  
 236 ELAVDGGMT 244

DB 236 ELAVDGGMT 244

RESULT 4  
 YFED\_BACSU STANDARD; PRT; 255 AA.  
 ID YFED\_BACSU  
 AC P39640;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical oxidoreductase yfwd (EC 1.-.-.-).  
 GN yfwd OR fpa-82D.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxId=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA MEDLINE=95020537; PubMed=7934828;  
 RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales M.,  
 RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,  
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,  
 RA Rapoport G., Danchin A.;  
 RA "Bacillus subtilis genome project: cloning and sequencing of the 97  
 kb region from 325 degrees to 333 degrees.";  
 RT Mol. Microbiol. 10:371-384(1993).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA MEDLINE=96044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriss R., Boursier L., Brans A., Braun M., Bridgnell S.C., Bron S.,  
 RA Broiliet S., Bruschi C.V., Caldwell B., Capano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings S.D., Daniel R.A.,  
 RA Denot F., Deville K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.A.,  
 RA Entlan K.D., Erlington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galazzi A., Galleron N.,  
 RA Ghm S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,  
 RA Guisepel G., Guy B.J., Haga K., Halech J., Harwood C.R., Heaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,  
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krog S., Kumano M.,  
 RA Kurita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porroli G., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche A., Roche B., Rose M., Sadleir Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,  
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weltzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis.";  
 RL Nature 390:249-256(1997).  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X73124; CAAS1638.1;  
 CC EMBL: 299123; CAB15799.1;  
 CC PIR: S39737; S39737.  
 CC HSSP: P50162; IAEI.  
 CC Subtilist; BG10628; yfwd.  
 CC InterPro: IPR002198; ADH\_short.  
 CC Pfam: PF00106; adh\_short; 1.  
 CC PRINTS: PR00080; SDRFAMILY.  
 CC PROSITE: P500061; ADH\_SHORT; 1.  
 CC NP\_BIND 11 33  
 CC Hypothetical protein; Oxidoreductase; Complete proteome.  
 CC FT ACT\_SITE 154 154  
 CC BY SIMILARITY.  
 CC SEQUENCE 255 AA; 27324 MW; 20AA2259BF88C9B CRC64;

Query Match  
 Best Local Similarity 31.0%; Score 400.5; DB 1; Length 255;  
 Matches 96; Conservative 45; Mismatches 107; Indels 11; Gaps 6;

DB 1 MSNRDQKVAITGGTGLGIAATKFEVEGAKVMTDRHSVDGEKAASVGTPOIOFROH 60  
 1 MINNLDKTYLVITGAGAGIGYAAVOAFLGQANVAVADIDEGAGAMRKNN--DRHIV 59  
 61 QHDSDEDGWTKLDATEKAFGPVSTLVNAGIAVANKSVEETTAEMRKLLAVNLGCVFET 120  
 60 QDTIDEAAGQHAHVESVHFGGLDVLLNAGIEIVAPIHHEWELSDMNKVLQVNLGMFL 119  
 121 GTRLGIOKMKKGGASTINNSIEGFPVDSGLAVNASKGAVRIMSKAALCALKDQYVR 180  
 120 MSKHALKHLMLAAGKG-NIINTCSVGLVAMPDIPAVNASKGGLVLTSMADVA--KIQ 176  
 181 VRVNTVHPGYIKTFLVDDLPAGEAASQRTKTPMGHIGEPNDIAVICYVLASN 233  
 177 IIRNCVCPGLIDPPLNKSFLNNEGTLIEIKKERAKYNPLRLKPEIATVAMPLASD 236  
 234 ESKFAFGSEFVVDGQYTAQ 252  
 237 LSSYMTGSATYADGQYTAQ 255

RESULT 5  
 Y019\_THEMEA STANDARD; PRT; 236 AA.  
 ID Y019\_THEMEA  
 AC 056318;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative oxidoreductase TM0019 (EC 1.-.-.-).  
GN TM0019.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;  
OC Thermotogaceae; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB / DSM 3109;  
RA MEDLINE=96125254; PubMed=8550425;  
RT Kletzin A., Adams M.;  
RT Molecular and phylogenetic characterization of pyruvate and 2-  
RT ketosulfolactate ferredoxin oxidoreductases from *Pyrococcus furiosus*  
RT and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*.  
RL J. Bacteriol. 178:248-257(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB / DSM 3109;  
RA MEDLINE=99287316; PubMed=10360571;  
RX Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT \*Evidence for lateral gene transfer between Archaea and Bacteria from  
RT genome sequence of *Thermotoga maritima*.  
RL Nature 393:323-329(1999).  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
DR EMBL: X85171; CA59459.1;  
DR EMBL: AE001690; AAD3113.1;  
DR HSSP: 070351; 1E6W.  
DR TIGR: TM0019;  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short.1.  
DR PRINTS: PR00080; SDRFAMILY.  
DR PROSITE: PS00061; ADH\_SHORT; 1.  
KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.  
FT NP\_BIND 9 33 NADP (BY SIMILARITY).  
FT ACT\_SITE 153 153 BY SIMILARITY.  
FT CONFLICT 130 136 KGGVIT -> TRMSDH (IN REF. 1).  
SQ SEQUENCE 236 AA: 28078 MW: D68160B1D7980C6B CRC64;  
Query Match 29.7% Score 383; DB 1; Length 256;  
Best Local Similarity 35.5% Pred. No. 2.9e-24;  
Matches 93; Conservative 48; Mismatches 95; Indels 26; Gaps 7;  
QY 5 LDGVAITGGLGILAIATKFEVGAQVMTDRHSDVG--EKAASVGTDPDIOFPO 61  
DB 2 LEGKVAVTGGGGGIGAIQAQFAENGKMKVIAIETDEAGVREMLERGL--DYTEVK 59  
QY 62 HDSSEDEMTLFDATETAFGPVSTLVNNAAGIAVANKSVETTTAEMRLLAVNLGVEFG 121  
DB 60 TDVADENSVMKMKVTRTVEIYGGVDVLVNNAAVMYSKIFERPLEFEMERIVRNLTGPIC 119  
QY 122 TRLGIORKNKGLASIIINMSIEGFVGPDLGAVNASKAGVRINMSKALDKADYD 181  
DB 120 SYCAEEMIKKG--GGVITINASTRAFOSEPTEPSYASKGGLVALATJHSLAV--SLSRVHI 176  
QY 182 RVNTVHPGYIKTPLVLDLFGAEAMSSORTKT-----PMGHIGPNPIAYICVILA 231  
DB 177 RVVISIPGMIEF-----SEMKKKSLAKKRPDLRPHDQHPAGVGNPDLIAHLCVELA 229

QY 232 SNE-SKEATGSEFVVDGTYAQ 252  
DB 230 DDEKAGFTGTGNFTVDCGMATK 251  
RESULT 6  
YD50\_MYCTU STANDARD; PRT; 247 AA.  
ID YD50\_MYCTU  
AC Q11020;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Putative oxidoreductase RV1350 (EC 1.-.-.-).  
GN FABG2 OR RV1350 OR MT1393 OR MYCT02B10.14.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
RT \*Deciphering the biology of *Mycobacterium tuberculosis* from the  
RT complete genome sequence.  
RL Nature 393:537-544(1998).  
CC [2]  
CC SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishal W.;  
RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and  
RT laboratory strains".  
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
DR EMBL: Z75555; CA9983.1;  
DR EMBL: AE007012; AAK45656.1;  
DR HSSP: 070351; 1E3S.  
DR TIGR: MT1393;  
DR TuberculList: RV1350;  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short.1.  
DR PRINTS: PR00080; SDRFAMILY.  
DR PROSITE: PS00061; ADH\_SHORT; 1.  
KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.  
FT NP\_BIND 9 34 NAD (BY SIMILARITY).  
FT ACT\_SITE 155 155 BY SIMILARITY.  
SQ SEQUENCE 247 AA: 25871 MW: 76CA07892E7BBA73 CRC64;  
Query Match 29.5% Score 380; DB 1; Length 247;  
Best Local Similarity 36.2% Pred. No. 4.8e-24;  
Matches 92; Conservative 52; Mismatches 94; Indels 16; Gaps 4;



```

OY 1 MSNRDGLKVAIITGTLGIGLAIAATKFEVEGAKVMTIDRHSVDGEKA-----AKSVGTP 60
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MASLNATTAITGGAOGIGLAIGFVAGARVVGADVLEATEVAAKRLGGDDVALAV 60
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 61 OHDSDEGCKTLPDTEKAFGPVSTLVNAGIAVNVKSVEETTAERKLLAVNLDGVEF 120
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 KCDVTQADVDILITFAVERFGLDVNVNAGITRDATMTTTEEOFDVIAVHLKGTWN 120
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 121 GTRLGIOPMKNGKGLASIIIMSSIEGFVGDPSLGAYNASKAVRINSKSAALDCLD 180
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 GTRILAAIIMREKRGKA-IVNMSSVSGKVMVGOTNYSAAKAGIVGTMKAKE--LAHLG 177
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 181 VRVNTVHGCIKTPVLDLPGEAEMSOR-----TKTPMGHIGEPNDIAYICVYLASN 234
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 IRVNAIAGLRASMT-----EAMPORIMQOKLAEVPMRAGESEVASAVFLASDL 230
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 235 SKFATGSEFVVDG 248
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 SSMYTGIVLDVTGG 244
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 7

DHG1\_BACME STANDARD: PRT: 261 AA.

```

AC P39482;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glucose 1-dehydrogenase I (EC 1.1.1.47) (GLCOH-1).
GN GDI1.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-IM 1030;
RA Mitamura T., Ebora R.V., Nakai T., Makino Y., Negoro S., Urabe I.,
RA Okada H.;
RT "Structure of isozyme genes of glucose dehydrogenase from Bacillus
RT megaterium IM1030."
RL J. Ferment. Bioeng. 70:363-369(1990).
CC -1- FUNCTION: MAY PLAY SOME ROLE IN SPORE GERMINATION.
CC -1- CATALYTIC ACTIVITY: Beta-D-glucose + NAD(P)(+) -> D-glucose-1,5-
CC lactone + NAD(P)H.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPORE germination.
CC -1- MISCELLANEOUS: PREFERS NADP TO NAD.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; D90043; BAA14099.1; -
DB PIR; J50385; J50385.
DB HSSP; P50163; 2AE1.
DB InterPro; IPR002198; ADH_short.
DB Pfam; PF00106; adh_short.1.
DB PRINTS; PR00080; SDRFAMILY.
DB PROSITE; PS00061; ADH_SHORT.1.
KW Oxidoreductase; NADP; Multigene family; Sporulation; Germination.
FT NP_BIND 11 35 NADP (BY SIMILARITY).
FT ACT_SITE 158 158
FT SEQUENCE 261 AA; 28117 MW; B92B72B89A160335 CRC64;

```

Query Match 28.8%; Score 371.5; DB 1; Length 261;  
 Best Local Similarity 34.2%; Pred. No. 2, 5e-23;  
 Matches 88; Conservative 51; Mismatches 105; Indels 13; Gaps 5;

```

OY 1 MSNRDGLKVAIITGTLGIGLAIAATKFEVEGAKVMTIDRHSVDGEKA-----AKSVGTP 54
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MYKDEGKVVITGSGTGKAMAIKFAFEKAKVNVNYSKE--EENSVLEEKIKV-- 56
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 55 DOIQFQHDSDEGCKTLPDTEKAFGPVSTLVNAGIAVNVKSVEETTAERKLLAVN 114
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 57 GEAIKVGKGVVAVESDVINLVQSSIKKEFGKLDVINNAGHENPNVSSHMSISDNKVI 116
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 115 LDGFFGTRLGIOPMKNGKGLASIIIMSSIEGFVGDPSLGAYNASKAVRINSKSAALD 174
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 117 LTGAFLGSRKAIKRYENDIKGTIVNMSSVHEKIPPLVHYVAKSGKMLMETLLEY 176
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 175 ALKDYDVRVNTVHGCIKTPVLDLPGEAEMSOR-----TKTPMGHIGEPNDIAYICVYLASN 233
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 APK-GIRVNNIGPGALNTPINAEKADPEQADVESMIPMGVIGPEEETIAVAAVLASS 234
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 234 ESKFATGSEFVVDG 250
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 235 EASYVTGITLPADGGMT 251
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 8

DHG1\_BACME STANDARD: PRT: 261 AA.

```

AC P40288;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glucose 1-dehydrogenase I (EC 1.1.1.47).
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF I-29, AND MUTAGENESIS.
RA STRAIN-IMG3;
RA MEDLINE=89197943; PubMed=2495285;
RA Makino Y., Negoro S., Urabe I., Okada H.;
RT "Stability-increasing mutants of glucose dehydrogenase from Bacillus
RT megaterium IMG3."
RL J. Biol. Chem. 264:6381-6385(1989).
CC [2]
CC REVISIONS.
CC Urabe I.;
CC Submitted (MAR-1989) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Beta-D-glucose + NAD(P)(+) -> D-glucose-1,5-
CC lactone + NAD(P)H.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPORE germination.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; J04805; AAA22475.1; -
DB PIR; A35528; A35528.
DB HSSP; P50163; 2AE1.
DB InterPro; IPR002198; ADH_short.
DB Pfam; PF00106; adh_short.1.
DB PRINTS; PR00080; SDRFAMILY.
DB PROSITE; PS00061; ADH_SHORT.1.
KW Oxidoreductase; NADP; Multigene family; Sporulation.
FT NP_BIND 11 35 NADP (BY SIMILARITY).
FT ACT_SITE 158 158
FT MUTAGEN 96 96 E->A,G,K; HEAT STABLE.
FT MUTAGEN 108 108 D->N; HEAT STABLE.
FT MUTAGEN 112 112 V->A; HEAT STABLE.

```

FT MUTAGEN 133 133 E->K: HEAT STABLE.  
 FT MUTAGEN 163 163 V->I: HEAT STABLE.  
 FT MUTAGEN 194 194 P->Q: HEAT STABLE.  
 FT MUTAGEN 210 210 E->K: HEAT STABLE.  
 FT MUTAGEN 217 217 Y->H: HEAT STABLE.  
 FT MUTAGEN 252 252 O->L: HEAT STABLE.  
 FT MUTAGEN 253 253 Y->C: HEAT STABLE.  
 FT MUTAGEN 258 258 A->G: HEAT STABLE.  
 SQ SEQUENCE 261 AA: 28085 MW: C23AC98D304EEB2F CRC64;

Query Match 28.7% Score 370.5; DB 1; Length 261;  
 Best Local Similarity 33.6%; Pred. No. 3.1e-23;  
 Matches 86; Conservative 51; Mismatches 108; Indels 11; Gaps 4;

OY 1 MSNRDGVKVAITGTGTLGTLATATKFEVEGAKVITP-----DRHSDVGEKAASVGPDP 55  
 DB 1 MYKDEGKVVITGSSTGLGSKMAIRFATKAKVYVVRKSKEDANSYLEIKKVGGEAT 60  
 OY 56 QIQFQHOSSDEGDKTLPDATEKAFGPVSTLVNAGIAVNSVEETTTAEKRLLAVNL 115  
 DB 61 AV---KGVTVESDVINTVQSAIKKEGKLDVMINNAGLEPNVSSHEKSLDMNKVDTIOTNL 117  
 OY 116 DQVFGTRIGRMKNKKGASIIINNSIEGFVGDPSLGAVNSKAVRINSKSALDCA 175  
 DB 118 TGAFLGSRBAIVFVENDIKGVINNSVHEKIPWPLEVHYAASKGKMLTETLAEYA 177  
 OY 176 LKDYDVAVTTPHGYIKTLP-VDDLPGAEEAMSGRTKTPMGHIGEPNDIAYICVYIASNE 234  
 DB 178 PK--GIRVNNIGPGAINPINKAFDPQADRVESMIPMGVIGEPETIAVAAMALASSE 235  
 OY 235 SKFAGSEFVVDGCT 250  
 DB 236 ASYVTGITLFLADGCMT 251

## RESULT 9

3BHD\_COMTE ID 3BHD\_COMTE STANDARD: PRT: 253 AA.

AC P19871;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE 3-beta-hydroxysteroid dehydrogenase (EC 1.1.1.51).  
 OC Comamonas testosteroni (Pseudomonas testosteroni).  
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.  
 OX NCBI\_TaxID=285;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 11996;  
 RX MEDLINE-93176721; PubMed-8382516;  
 RA Abalain J.H., di Stefano S., Amet Y., Quemener E.,  
 RA Abalain-Colloc M.L., Floch H.H.;  
 RT Cloning, DNA sequencing and expression of (3-17)beta hydroxysteroid  
 RT dehydrogenase from Pseudomonas testosteroni.  
 RL J. Steroid Biochem. Mol. Biol. 44:133-139(1993).  
 RN [2]  
 RP SEQUENCE:  
 RX MEDLINE-91224127; PubMed-2026158;  
 RX Yin S.-J., Vagelopoulos N., Lundquist G., Joernvall H.;  
 RT Pseudomonas 3 beta-hydroxysteroid dehydrogenase. Primary structure  
 RT and relationships to other steroid dehydrogenases.  
 RL Eur. J. Biochem. 197:359-365(1991).  
 RN [3]  
 RP CRYSTALLIZATION.  
 RC STRAIN-ATCC 11996;  
 RX MEDLINE-96184893; PubMed-8617258;  
 RA Benach J., Knapp S., Oppermann U.C.T., Haeglund O., Joernvall H.,  
 RA Ladenstein R.;  
 RT Crystallization and crystal packing of recombinant 3 (or 17) beta-  
 RT hydroxysteroid dehydrogenase from Comamonas testosteroni ATCC  
 RT 11996.  
 RL Eur. J. Biochem. 236:144-148(1996).  
 CC -1- CATALYTIC ACTIVITY: Testosterone + NAD(P)(+) = androst-4-ene-3,17-

CC  
 CC  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 CC -1- DATABASE: NAME-Worthington enzyme manual;  
 CC WWW="http://www.worthington-biochem.com/manual/H/STDH.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sdb.cn/announce/](http://www.isb-sdb.cn/announce/or send an email to license@isb-sdb.cn)  
 CC or send an email to [license@isb-sdb.cn](mailto:license@isb-sdb.cn)).  
 CC  
 CC EMBL: X6379; CAA44977.1;  
 DR PIR: S15390; S15390.  
 DR HSSP: P19992; 1HDC.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT; 1.  
 DR OXidoreductase; NAD; Steroid metabolism.  
 DR INIT. MET 0  
 FT NP\_BIND 10 34 NAD (BY SIMILARITY).  
 FT ACT\_SITE 150 150  
 FT CONFLICT 13 14 VV -> GG (IN REF. 2).  
 FT CONFLICT 39 39 N -> NE (IN REF. 2).  
 FT CONFLICT 177 177 MISSING (IN REF. 2).  
 FT CONFLICT 240 240 S -> G (IN REF. 2).  
 SQ SEQUENCE 253 AA: 26932 MW: CE677D9021458B39 CRC64;

Query Match 28.7% Score 370.5; DB 1; Length 253;  
 Best Local Similarity 35.6%; Pred. No. 3.2e-23;  
 Matches 89; Conservative 49; Mismatches 102; Indels 10; Gaps 6;

OY 2 SNRDLGKVAITGTGTLGTLATATKFEVEGAKVITPDRHSDVGEKAASVGPDPDIOFQ 61  
 DB 1 TNRLOGKVALVTVYASGVGLGVKLLGEGAKVAFSDINA-AGQOLAELG--ERSMFR 57  
 OY 62 HDSDEGDKTLPDATEKAFGPVSTLVNAGIAVNSVEETTTAEKRLLAVNLGDFVFG 121  
 DB 58 HGVSSPADMTLVAAVQRRIGTLNVLNAGIILLPGMETGRLEDFSLINTEFSVIG 117  
 OY 122 TRLGORMKNKKGASIIINNSIEGFVGDPSLGAVNSKAVRINSKSALDCAKDIDV 181  
 DB 118 COOGIAAME--TGGSIIINNASVSMPLIPQYAGYSKAAVSALTRAAALSCROGYAI 175  
 OY 182 -RVNTVHPGYIKTLPVD-DLP--GAEEAMSGRTKTPMGHIGEPNDIAYICVYIASNEK 236  
 DB 176 RRVNSIHPPGCIITPMQASLPKGVSKEMVLDKRLNRAVYMPERIAQLVFLASDESS 235  
 OY 237 FATGSEFVVD 246  
 DB 236 VMSGSELHAD 245

## RESULT 10

FABG\_THEME ID FABG\_THEME STANDARD: PRT: 246 AA.

AC Q9X248;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-  
 DE acyl carrier protein reductase).  
 GN FABG OR TM1724.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogales;  
 OC Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=2336;  
 RN [1]  
 RP SEQUENCE FROM N.A.





\*The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 RT sublin1.1: Nature 390:249-256(1997).  
 RL -1- CATALYTIC ACTIVITY: Beta-D-glucose + NAD(P)(+) -> D-glucono-1,5-lactone + NAD(P)H.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- INDUCTRON: IT IS INDUCED AT STAGE III OF THE SPOULATION.  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL: M12276; AAA22463.1; -  
 DR EMBL: D50453; BAA09024.1; -  
 DR EMBL: 299106; CAB12201.1; -  
 DR PIR: S36090; S36090.  
 DR HSSP: P50162; IAE1.  
 DR Subtilist: BG10545; gdh.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short.1.  
 DR PRINTS: PRO0080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT.1.  
 DR Oxidoreductase: NADP: Sporulation; Complete proteome.  
 FT NP\_BIND 11 35 NADP (BY SIMILARITY).  
 FT ACT\_SITE 158 158 BY SIMILARITY.  
 FT CONFLICT 148 150 EVI -> AF (IN REF. 1).  
 FT SEQUENCE 261 AA; 28090 MW; 5894C17DB8F14965 CRC64;

Query Match 27.8%; Score 358.5; DB 1; Length 261;  
 Best Local Similarity 33.9%; Pred. No. 2.9e-22;  
 Matches 87; Conservative 45; Mismatches 104; Indels 21; Gaps 5;

QY 5 LDGKVAITGGTGLGIGLAIAATKFEVEEGAKVMIT-----DRHSDVGEKAKSVGTPDOIOF 59  
 DB 5 LKGVAAITGAASGLGKAMAFRGKBOAKVINYNSKQDNEVEKKEVYKAGC---ENVV 61  
 QY 60 FOHDSDEGCKWTLKFDATERAFGVPSTLVNNAAGAVNKSVEETTTAERKLLAVNLGDF 119  
 DB 62 VOGDVTKKEEDVKNIVYOTAIKEFGTLDIMINNAAGLENVPSPHEMPLKDDKVTGNTLGA 121  
 QY 120 FCTRLGIGQMKNGKGLASIIINMSSIEGFVGDPSIGAYNASKGAVRINMSKSAALDCAIKDY 179  
 DB 122 IGSREAIKIFVENDIKGNVIMSSVHEVTPLEFVHYAASGGIKLMTETLALVAPK-- 179  
 QY 180 DVAVNTVHVGITKTPLVDDLPGAEAMSSQRTKT-----PMGHIGEPNDIAYICVYLASN 233  
 DB 180 GIRVNNIGGAIINTPI-----NAEKFADPKQKADVESMIPGVIYIGEPPEIAVAAMWLASK 234  
 QY 234 ESKFATGSEFVVDGGYT 250  
 DB 235 EASYVTGITLFLADGGMT 251

RESULT 14  
 DHG4\_BACME  
 ID DHG4\_BACME STANDARD; PRT; 261 AA.  
 AC P39485;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Glucose 1-dehydrogenase IV (EC 1.1.1.47) (GLCDH-IV).  
 GN GDBIT.  
 OS *Bacillus megaterium*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1404;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-1AM 1030;  
 RX MEDLINE-92332436; PubMed-1629157;  
 RA Nagao T., Mitamura T., Wang X.H., Negoro S., Yomo T., Urabe I.,  
 RA Okada H.;  
 RT "Cloning, nucleotide sequences, and enzymatic properties of glucose  
 RT dehydrogenase isozymes from *Bacillus megaterium* IAM1030.";  
 RL J. Bacteriol. 174:5013-5020(1992).  
 CC -1- CATALYTIC ACTIVITY: Beta-D-glucose + NAD(P)(+) -> D-glucono-1,5-lactone + NAD(P)H.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- MISCELLANEOUS: PREFERRED NAD TO NADP; 2M NaCl ENHANCES ITS PH AND THERMOSTABILITY.  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL: D10626; BAA01476.1; -  
 DR HSSP: P50162; IAE1.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short.1.  
 DR PRINTS: PRO0080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT.1.  
 DR Oxidoreductase: NAD; Multiligene family.  
 FT NP\_BIND 11 35 NAD (BY SIMILARITY).  
 FT ACT\_SITE 158 158 BY SIMILARITY.  
 FT SEQUENCE 261 AA; 28157 MW; 6FBEC9397BCF17C CRC64;

Query Match 27.7%; Score 357.5; DB 1; Length 261;  
 Best Local Similarity 32.1%; Pred. No. 3.5e-22;  
 Matches 85; Conservative 51; Mismatches 100; Indels 29; Gaps 6;

QY 1 MSNLDGKVAITGGTGLGIGLAIAATKFEVEEGAKVMIT-----DRHSDVGEKAKSV 51  
 DB 1 MYDILAKRVVITGSGTGLRAMAVRFGQEKAVVINYNEBALDAKKEVEAGGQAT 60  
 QY 52 GTPDOIOFPOHDSDEGCKWTLKFDATERAFGVPSTLVNNAAGAVNKSVEETTTAERKLL 111  
 DB 61 -----YOGDVTKKEEDVKNIVYOTAIKEFGTLDIMINNAAGLENVPSPHEMPLKDDKVTGNTLGA 113  
 QY 112 AVNLGDFGFGTGLGIGQMKNGKGLASIIINMSSIEGFVGDPSIGAYNASKGAVRINMSKSA 171  
 DB 114 DTNLTGAFILGSREAIKIFVENDIKGNVIMSSVHEMTPLEFVHYAASGGIKLMTETL 173  
 QY 172 LDCALKRDYDVAVNTVHVGITKTPLVDDLPGAE- EAMSQRTKTPMGHIGEPNDIAY 225  
 DB 174 LETAAPK--GIRVNNIGGAIINTPI-----NAEKFADPKQKADVESMIPGVIYIGEPPEIAVA 226  
 QY 226 ICVYLANSKESKATGSEFVVDGGYT 250  
 DB 227 VAAFLASSQASVYTGITLFLADGGMT 251

RESULT 15  
 DHG2\_BACME  
 ID DHG2\_BACME STANDARD; PRT; 261 AA.  
 AC P39483;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Glucose 1-dehydrogenase II (EC 1.1.1.47) (GLCDH-II).  
 GN GDBIT.  
 OS *Bacillus megaterium*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1404;  
 RN [1]  
 RP SEQUENCE FROM N.A.



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 17:13:46 ; Search time 35 seconds  
(without alignments)  
1483.539 Million cell updates/sec

Title: US-09-910-033A-2  
Perfect score: 1290  
Sequence: 1 MSNRLDGKVAIIITGTLGIG.....NESKFAFGSEFVVDGCTAAQ 252

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	36.4	261	16 09AB11	09ab11 caulobacter
2	452.5	35.1	250	2 0937L4	0937L4 comamonas t
3	438	34.0	254	16 09K4G9	09K4G9 streptomyc
4	427.5	33.1	244	2 08VL53	08VL53 xanthobacte
5	411	31.9	251	2 09RNK5	09RNK5 zymomonas m
6	411	31.9	254	2 052587	052587 comamonas t
7	395.5	30.7	256	2 0930G5	0930G5 brevibacter
8	391	30.3	249	2 056841	056841 xanthobacte
9	388	30.1	260	17 08U3B3	08U3B3 pyrococcus
10	386.5	30.0	247	2 093RMO	093RMO acinetobact
11	386.5	30.0	251	2 09F7E0	09F7E0 acinetobact
12	378.5	29.3	258	2 09JN17	09JN17 agrobacteri
13	376.5	29.2	255	16 098ER0	098ER0 rhizobium 1
14	374.5	29.0	248	16 031680	031680 bacillus su
15	374.5	29.0	250	16 098C63	098C63 rhizobium 1
16	373	28.9	249	2 09LBG5	09LBG5 geobacillus

17	372.5	28.9	253	16 0930L7	0930L7 rhizobium m
18	371.5	28.8	252	16 08XWV9	08XWV9 ralsstonia s
19	370	28.7	246	2 09EX74	09EX74 rhodococcus
20	369.5	28.6	251	16 092PY8	092PY8 rhizobium m
21	368	28.5	253	16 09CH41	09CH41 lactococcus
22	367	28.4	258	16 033339	033339 mycobacteri
23	366	28.4	272	16 099RG1	099RG1 staphylococ
24	364.5	28.3	255	16 P95286	P95286 mycobacteri
25	364	28.2	264	16 09AK65	09AK65 streptomyc
26	360.5	27.9	251	16 08RG25	08RG25 fusobacteri
27	360	27.9	251	2 09FDK2	09FDK2 zymomonas m
28	359	27.8	256	16 0982N5	0982N5 rhizobium 1
29	359	27.8	269	16 08UKD8	08UKD8 agrobacteri
30	358	27.8	270	6 09WYP6	09WYP6 bos taurus
31	357.5	27.7	251	16 08UKA9	08UKA9 agrobacteri
32	357.5	27.7	258	16 0930L5	0930L5 rhizobium m
33	357.5	27.7	261	2 09F5L5	09F5L5 bacillus su
34	357.5	27.7	271	16 08ZGK4	08ZGK4 yersinia pe
35	357	27.6	296	16 09ABX6	09ABX6 caulobacter
36	356.5	27.6	248	16 08U9B5	08U9B5 agrobacteri
37	356.5	27.6	256	16 092MR3	092MR3 rhizobium m
38	353.5	27.4	255	16 08R0C3	08R0C3 thermotoga
39	351	27.2	286	16 091IX3	091IX3 pseudomonas
40	350	27.1	257	16 09WYD3	09WYD3 thermotoga
41	348.5	27.0	255	16 092RW7	092RW7 rhizobium m
42	347.5	26.9	268	17 09HLN6	09HLN6 thermoplasm
43	347	26.9	243	16 09K0C9	09K0C9 bacillus ha
44	346.5	26.9	254	2 09RH24	09RH24 zymomonas m
45	346.5	26.9	255	16 09WYS2	09WYS2 thermotoga

ALIGNMENTS

RESULT 1

09AB11 ID 09AB11 PRELIMINARY: PRT: 261 AA.

AC 09AB11:

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase.

GN CC0246.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.

OX NCBI\_TaxID=155892;

RM [1]

RP SEQUENCE FROM N.A. / GB15;

RC STRAIN=ATCC 19089 / GB15;

RX MEDLINE=21173698; PubMed=11259647;

RA Niedman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Holt D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Uterback T., Tran K., Wolf A., Yamshchikov J., Ermolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,

RT "Complete genome sequence of Caulobacter crescentus."

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC EMBL: AE005698; AKR2233.1; -

DR HSSP: P19992; IHNC.

DR	TIGR: CC0246; -
DR	InterPro: IPR002198; ADH\_short.
DR	Pfam: PF00106; adh\_short; 1.
DR	PRINTS: PR00080; SDRFAMILY.
DR	PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.
KW	Oxidoreductase; Complete proteome.
SEQ	SEQUENCE 261 AA: 27011 MW; CE011D872D4EBA6D CRC64;
Query Match	36.4%; Score 470; DB 16; Length 261;

Best Local Similarity 40.7%; Pred. No. 2.9e-27;  
Matches 107; Conservative 46; Mismatches 90; Indels 20; Gaps 6;

QY 2 SNRLDGKVAIIITGTLGIGLAITATKFEVGEAKWITDRHSVGEKAASVGTPOIQ--- 58  
DB 5 TGRVAGKKAFFITGAGGAGCAGAAAGMLAKESAKVAL---ADINLAGAOVA--DEINAAH 58  
QY 59 -----FEOHSSDEDEGWTKLFEDATEKAFGVPSTLVNAGIAVNSVEETTAEMRKLLA 112  
DB 59 GAGTAFAPFELDYTOEDQWIDVLEKATAAGGLSVLVNAGIGCGPIESLDIFGLMKVMS 118  
QY 113 VLDGVEFETRLGIGRMKKGAGSIINNSIGFVCDPBLGAVNASKGAVRMSKAL 172  
DB 119 VAVDSVFELAKAHLTHMRHOPe-STIINSSINGLANGSPAYNASKAAMVLLSKNIAL 177  
QY 173 DCAIKDYRVNTHVPGYIKTPLYDDLP---GAEAMSORTK-TPMGHIGEPNDIAYICV 228  
DB 178 YCAKKMLDRIKNSIHFTIDTPIIDGFSARFKGEBAFAKLARQVPLGRIGEPNDIANNVL 237  
QY 229 YLASNESKATGSEFVVDGYYA 251  
DB 238 YLASDESKFMTGAETIKVDGISA 260

## RESULT 2

0937L4

PRELIMINARY; PRT: 250 AA.

AC 0937L4  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
DE Cyclohexanol dehydrogenase (EC 1.1.1.1).  
GN CPMB.  
OS Comamonas testosteroni (Pseudomonas testosteroni).  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.  
RN NCBI\_TaxID=285;  
RP SEQUENCE FROM N.A.  
RC STRAIN-NCIMB 9872;  
RA van Bellen J.B., Fritzsche U., Seeger M., Smits T.H.M., Wiltolt B.;  
RT "Cloning of Baeyer-Villiger monooxygenases from Comamonas,  
xanthobacter and Rhodococcus via PCR with highly degenerate primers."  
RT Submitted (NOV-2001) to the EMBL/Genbank/DDI databases  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.  
DR EMBL: AJ418060; CAD10799.1;  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
KM Oxidoreductase.  
SQ SEQUENCE 250 AA: 26626 MW: 7642303CBAEBB858E CRC64;

Query Match 35.1%; Score 452.5; DB 2: Length 250;  
Best Local Similarity 38.8%; Pred. No. 5.4e-26;  
Matches 97; Conservative 48; Mismatches 100; Indels 5; Gaps 4;

QY 4 RLDDGVAIIITGTLGIGLAITATKFEVGEAKWITDRHSVGEKAASVGTPOIQFQH- 62  
DB 3 RVNKKVVLVTGAGMGGLTCTLLAREGATVYLSDMNEETGHOVAVEIRQCGKAHFLH 62  
QY 63 DSSDEGWTKLFEDATEKAFGVPSTLVNAGIAVNSVEETTAEMRKLLAVNLDGVEFET 122  
DB 63 DVTNENHTGAVDTILAESDRIDLAVNNAGILTLKPVQDTSNEEMWRIETINRSVFLGT 122  
QY 123 RLGIORMKNGKAGSIINNSIEGFVCDPBLGAVNASKGAVRMSKALDCAIKDYDV 182  
DB 123 RAVLEPMKAKHGC-IVNVSITGLVAGAGAAAYESKAVRLFTACAVD--LAPNIR 179  
QY 183 VNVTHPGYIKTPLYDD-LPGAEAMSORTKTPMGHIGEPNDIAYICVYLASNESKATGS 241  
DB 180 VNSVHPGVIATPMTQOILDAPOSARALLGLPTLLGRAAQPMVEISOAVLFLVSDASFVHGS 239  
QY 242 EFVVDGYYA 251  
I | | | | |

DB 240 ELVVDGYYA 249

## RESULT 3

09K4G9

PRELIMINARY; PRT: 254 AA.

AC 09K4G9  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
DE Putative oxidoreductase.  
GN SC04169 OR SC066.06.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OC NCBI\_TaxID=1902;  
RN NCBI  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RL Mol. Microbiol. 21:77-96(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2) / M145;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Horsby T., Howard S.,  
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2)."  
RL Nature 417:141-147(2002).  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.  
DR EMBL: AL358692; CAB94073.1;  
DR HSSP: P19992; IHDC.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR PRINTS: PR00080; SDRFAMILY.  
DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.  
KM Oxidoreductase.  
SQ SEQUENCE 254 AA: 26132 MW: A372F691F1B8672 CRC64;

Query Match 34.0%; Score 438; DB 16: Length 254;  
Best Local Similarity 40.7%; Pred. No. 6.6e-25;  
Matches 100; Conservative 40; Mismatches 100; Indels 6; Gaps 4;

QY 5 LDGVAIIITGTLGIGLAITATKFEVGEAKWITDRHSVGEKAASVGTPOIQFQHDS 64  
DB 4 LTGKNVVIITGARGGAGAEARQAVAGAVLITVDDDOENAREIG--DRARFLHHVD 61  
QY 65 SDEGWTKLFEDATEKAFGVPSTLVNAGIAVNSVEETTAEMRKLLAVNLDGVEFETRL 124  
DB 62 TSEDMRSRADEAVTEFGALHGLVNNAGISTGTPLESESVDHFRKLVNLTGVFIQMKT 121  
I | | | | |









Db 237 VTGSQVYVDGAYTSK 251

## RESULT 12

09JUN17

ID 09JUN17 PRELIMINARY: PRT: 258 AA.

AC 09JUN17; 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)

GN YHG.

OS Agrobacterium tumefaciens.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI\_TaxID=358;

RN [1]

RP MEDLINE-96236046; PubMed-8655509;

RX Kim K.S., Farrand S.K.;

RA "Ti plasmid-encoded genes responsible for catabolism of the crown gall

RT opine homopline by Agrobacterium tumefaciens are homologs of the T-

RT region genes responsible for synthesis of this opine by the plant

RT tumor."

RL J. Bacteriol. 178:3275-3284(1996).

RN [2]

RP MEDLINE-99141607; PubMed-9987134;

RX Lyl S.M., Jafri S., Minns S.C.;

RA "Mannoplinic acid and agropinolic acid catabolism region of the octopline-

RT type T1 plasmid pT119935."

RL Mol. Microbiol. 31:339-347(1999).

RN [3]

RP SEQUENCE FROM N.A.

RA Minns S.C., Zhu J., Oger P.M., Schrammeyer B., Hooykaas P.J.;

RA Farrand S.K.;

RT "Octopline-type T1 plasmid sequence."

RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

CC (SDR) FAMILY.

CC EMBL: AF242881; AAF77146.1; -

DR HSSP: P50162; IAEI.

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.

KW Oxidoreductase; Plasmid.

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 13

09BENO

ID 09BENO PRELIMINARY: PRT: 255 AA.

AC 09BENO; 01-OCT-2001 (Tremblrel. 18, Created)

DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)

DE 01-OCT-2001 (Tremblrel. 18, Last annotation update)

GN MRA4172.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 14

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 15

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 16

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 17

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 18

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 19

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

## RESULT 20

031680

ID 031680 PRELIMINARY: PRT: 248 AA.

AC 031680; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)

GN YKVO protein.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAFE30309;

RX MEDLINE-21082930; PubMed-11214968;

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.;

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.;

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.;

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50888.1; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short.1.

DR PRINTS: PR00080; SDRFAMILY.

KW Complete proteome.

SO SEQUENCE 255 AA; 26441 MW; 7EAF89BD46810117 CRC64;

Db 237 VTGSQVYVDGAYTSK 251

RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA MEDLINE-98044033; PubMed-9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Bourcier L., Brans A., Braun M., Brinell S.C., Bron S.,  
RA Broillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Chim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Nodack M.,  
RA Moore D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Prescan E., Pulic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadleir Y.,  
RA Sato T., Seanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tostato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambolt R., Wedler E., Wedler H., Weitenegger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
RT subtilis.";  
RT Nature 390:249-256(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,  
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASIS/REDUCTASES  
CC (SDR) FAMILY.  
CC EMBL: 299111; CAB13250.1; -  
DR HSP; P47227; 1BDB.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR Oxidoreductase; Complete proteome.  
SQ SEQUENCE 248 AA; 26245 MW; EACA490C1239524D CRC64;

Query Match 29.0%; Score 374.5; DB 16; Length 248;  
Best Local Similarity 37.5%; Pred. No. 3.3e-20;  
Matches 93; Conservative 40; Mismatches 106; Indels 9; Gaps 5;  
OY 4 RLDGKVAITGCTGIGLATAATKFEVGGAKVMTDRHSDVGEKAKSVGTPODIOFPHD 63  
DB 3 KREGKALVATGSGIGLATAOKFVNEGAVYITGRONEDKAVNOG--KNTGVGVD 60  
OY 64 SDEDEGWTLPDATERAFGVSTLVNAGIAVNSVEETTAEMRKLLAVNLGVGFGTR 123  
DB 61 ISKLELDLRIYKIKQEKGLDILFANAGIGNFLPGSEITBEQVDRFDIVKGIFFVQ 120  
OY 124 LGIQRKKNKGLGASIIINMSIEGFGVDPISGAVNASKGAVRIMSKAALDCAKDYDVR 183  
DB 121 KALSLFPDK--VGSIIIVTGSTAGISGNPAFVYGASKAALRALVRNMLD--LKGTETIRV 176  
OY 184 NTVHPGYIKTPLVDLPG--AEEAM--SQRTPMGHIGEPNDIYICVYLAISNESKPTG 240  
DB 177 NVVSPGILPAYDELFGDALEEVLENSRNTVPAGKVTPEEVANAVSFILASDESSYLTG 236  
OY 241 SEFVVDG 248  
DB 237 VELFVDG 244

RESULT 15  
O98C63 PRELIMINARY; PRT; 250 AA.  
ID O98C63;  
AC O98C63;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE Glucose dehydrogenase.  
GN MRF5280.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-WAFR30309;  
RX MEDLINE-21082930; PubMed-11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa K., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.,  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RT DNA Res. 7:331-338(2000).  
DR EMBL: AP003006; BAB51758.1; -  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR PRINTS: PR00080; SDRFAMILY.  
KW Complete proteome.  
SQ SEQUENCE 250 AA; 26407 MW; 1736D8C3F6711274 CRC64;

Query Match 29.0%; Score 374.5; DB 16; Length 250;  
Best Local Similarity 37.1%; Pred. No. 3.4e-20;  
Matches 93; Conservative 47; Mismatches 98; Indels 13; Gaps 5;  
OY 4 RLDGKVAITGCTGIGLATAATKFEVGGAKVMTDRHSDVGEKAKSVGTPODIOFPHD 63  
DB 3 KLEGKALVITGSGSGIGLATAKRFVEGCAHVITGRKEKELKEAAFI--MRNVTYVGD 60  
OY 64 SDEDEGWTLPDATERAFGVSTLVNAGIAVNSVEETTAEMRKLLAVNLGVGFGTR 123  
DB 61 VELLELDLRIYAVVKEKHGIDVLFANAGAGTIAFLAAATEAHFDQFVDVVKGLFFVQ 120  
OY 124 LGIQRKKNKGLGASIIINMSIEGFGVDPISGAVNASKGAVRIMSKAALDCAKDYDVR 183  
DB 121 KALPLFKD--GGSIIINSSVNLGLPGFSTYAAASAAVARNFSRAATLE--LKDRKIRV 175  
OY 184 NTVHPGYIKTPLVDLPG--AEEAM--SQRTPMGHIGEPNDIYICVYLAISNESKPTG 237  
DB 176 NMSPEALIEPLALETITGLTPBEQDAVAFQAFQIPGRGRKPEETIAAATFLASDDSSY 235  
OY 238 ATGSEFVVDG 248  
DB 236 VTGVDLAVDGG 246

Search completed: April 22, 2003, 17:16:27  
Job time : 38 secs